

TENT COOPERATION TRE Y

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

Commissioner
 US Department of Commerce
 United States Patent and Trademark
 Office, PCT
 2011 South Clark Place Room
 CP2/5C24
 Arlington, VA 22202
 ETATS-UNIS D'AMERIQUE
 in its capacity as elected Office

Date of mailing (day/month/year) 09 May 2001 (09.05.01)	
International application No. PCT/NZ00/00174	Applicant's or agent's file reference 18134/8X109
International filing date (day/month/year) 04 September 2000 (04.09.00)	Priority date (day/month/year) 02 September 1999 (02.09.99)
Applicant GLARE, Travis, Robert et al	

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International Preliminary Examining Authority on:
 26 March 2001 (26.03.01)

☐ in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was

☐ was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No.: (41-22) 740.14.35	Authorized officer Claudio Borton Telephone No.: (41-22) 338.83.38
---	--

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 18134/8X109	FOR FURTHER ACTION see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. PCT/NZ 00/ 00174	International filing date (day/month/year) 04/09/2000	(Earliest) Priority Date (day/month/year) 02/09/1999
Applicant NEW ZEALAND PASTORAL AGRICULTURE RESEARCH INSTI...		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 6 sheets.



It is also accompanied by a copy of each prior art document cited in this report.

1. Basis of the report

- a. With regard to the **language**, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.



the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

- b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :



contained in the international application in written form.



filed together with the international application in computer readable form.



furnished subsequently to this Authority in written form.



furnished subsequently to this Authority in computer readable form.



the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.



the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. ☒ **Certain claims were found unsearchable** (See Box I).

3. ☐ **Unity of invention is lacking** (see Box II).

4. With regard to the **title**,



the text is approved as submitted by the applicant.



the text has been established by this Authority to read as follows:

NUCLEOTIDES SEQUENCES ENCODING AN INSECTIDAL PROTEIN COMPLEX FROM SERRATIA

5. With regard to the **abstract**,



the text is approved as submitted by the applicant.



the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No.



as suggested by the applicant.



because the applicant failed to suggest a figure.



because this figure better characterizes the invention.



None of the figures.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 17

Present claim 17 relates to a ligand defined by reference to a desirable characteristic or property, namely binding to the polypeptide of claim 15.

The claim covers all products having this characteristic or property, whereas the application provides no support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for such products.

In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product by reference to a result to be achieved.

Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/NZ 00/00174

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/31 C12N15/70 C12N15/82 C07K14/24 C12Q1/68
A01N63/02 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K A01H C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

BIOSIS, EPO-Internal, STRAND, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JACKSON T A ET AL: "PATHOGEN TO PRODUCT DEVELOPMENT OF SERRATIA-ENTOMOPHILA ENTEROBACTERIACEAE AS A COMMERCIAL BIOLOGICAL CONTROL AGENT FOR NEW ZEALAND GRASS GRUB COSTELYTRA-ZEALANDICA" JACKSON, T. A. AND T. R. GLARE (ED.). USE OF PATHOGENS IN SCARAB PEST, 1992, pages 191-198, XP000997900 0-946707-35-9. 1992 the whole document --- -/--	32

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

° Special categories of cited documents:

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

Date of the actual completion of the international search

23 May 2001

Date of mailing of the international search report

06/06/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Holtorf, S

INTERNATIONAL SEARCH REPORT

International Application No

PCT/NZ 00/00174

C.(Continuation) DOCUMENTS CONSIDERED RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>GRKOVIC STEVE ET AL: "Genes Essential for Amber Disease in Grass Grubs Are Located on the Large Plasmid Found in Serratia entomophila and Serratia proteamaculans." APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 61, no. 6, 1995, pages 2218-2223, XP000994573 ISSN: 0099-2240 cited in the application the whole document</p> <p>---</p>	
A	<p>GLARE TRAVIS R ET AL: "Plasmid transfer among several members of the family Enterobacteriaceae increases the number of species capable of causing experimental amber disease in grass grub." FEMS MICROBIOLOGY LETTERS, vol. 139, no. 2-3, 1996, pages 117-120, XP000998482 ISSN: 0378-1097 cited in the application the whole document</p> <p>---</p>	
A	<p>WO 99 42589 A (NOVARTIS ERFIND VERWALT GMBH ;NOVARTIS AG (CH); KRAMER VANCE CARY) 26 August 1999 (1999-08-26) the whole document</p> <p>---</p>	
A	<p>WO 98 08932 A (DOW AGROSCIENCES LLC ;WISCONSIN ALUMNI RES FOUND (US)) 5 March 1998 (1998-03-05) the whole document</p> <p>---</p>	
A	<p>WO 98 08388 A (MORGAN JAMES ALUN WYNNE ;JARRETT PAUL (GB); ELLIS DEBORAH JUNE (GB) 5 March 1998 (1998-03-05) the whole document</p> <p>---</p>	
A	<p>WO 97 17432 A (WISCONSIN ALUMNI RES FOUND) 15 May 1997 (1997-05-15) the whole document</p> <p>---</p>	
A	<p>BOWEN D ET AL: "INSECTICIDAL TOXINS FROM THE BACTERIUM Photorhabdus liminescens" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, ,US, vol. 280, 26 June 1998 (1998-06-26), pages 2129-2132, XP002115650 ISSN: 0036-8075 cited in the application</p> <p>---</p> <p style="text-align: center;">-/--</p>	

INTERNATIONAL SEARCH REPORT

International Application No

PCT/NZ 00/00174

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NUNEZ-VALDEZ M E ET AL: "The amb2 locus from <i>Serratia entomophila</i> confers anti-feeding effect on larvae of <i>Costelytra zealandica</i> (Coleoptera: Scarabaeidae)" GENE,NL,ELSEVIER BIOMEDICAL PRESS. AMSTERDAM, vol. 172, no. 1, 12 June 1996 (1996-06-12), pages 75-79, XP004042712 ISSN: 0378-1119 cited in the application -----	
P,X	HURST MARK R H ET AL: "Plasmid-located pathogenicity determinants of <i>Serratia entomophila</i> , the causal agent of amber disease of grass grub, show similarity to the insecticidal toxins of <i>Photobacterium luminescens</i> ." JOURNAL OF BACTERIOLOGY, vol. 182, no. 18, September 2000 (2000-09), pages 5127-5138, XP002166799 ISSN: 0021-9193 the whole document -----	1-4, 9-16, 21-27, 31,41

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/AZ 00/00174

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9942589 A	26-08-1999	AU 3028699 A EP 1054972 A	06-09-1999 29-11-2000
WO 9808932 A	05-03-1998	AU 729228 B AU 1050997 A AU 2829997 A BR 9606889 A BR 9711441 A CA 2209659 A EP 0797659 A EP 0970185 A HU 9900768 A JP 2000515024 T PL 321212 A PL 332033 A SK 24699 A SK 93197 A TR 9901126 T WO 9717432 A	25-01-2001 29-05-1997 19-03-1998 28-10-1997 24-10-2000 15-05-1997 01-10-1997 12-01-2000 28-06-1999 14-11-2000 24-11-1997 16-08-1999 10-04-2000 06-05-1998 21-07-1999 15-05-1997
WO 9808388 A	05-03-1998	AU 4024997 A BR 9711285 A CN 1233938 A EP 0923295 A TR 9900435 T ZA 9707373 A	19-03-1998 17-08-1999 03-11-1999 23-06-1999 21-06-1999 15-02-1999
WO 9717432 A	15-05-1997	AU 729228 B AU 1050997 A BR 9606889 A CA 2209659 A EP 0797659 A HU 9900768 A PL 321212 A PL 332033 A SK 93197 A AU 2829997 A BR 9711441 A EP 0970185 A JP 2000515024 T SK 24699 A TR 9901126 T WO 9808932 A	25-01-2001 29-05-1997 28-10-1997 15-05-1997 01-10-1997 28-06-1999 24-11-1997 16-08-1999 06-05-1998 19-03-1998 24-10-2000 12-01-2000 14-11-2000 10-04-2000 21-07-1999 05-03-1998

PATENT COOPERATION TREATY

From the:
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To: JAMES & WELLS Private Bag 3140 HAMILTON New Zealand		PCT NOTIFICATION OF TRANSMITTAL OF INTERNATIONAL PRELIMINARY EXAMINATION REPORT (PCT Rule 71.1)	
		Date of mailing <i>day/month/year</i> 1 DEC 2001	
Applicant's or agent's file reference 18134/8X109SB		IMPORTANT NOTIFICATION	
International Application No. PCT/NZ00/00174	International Filing Date 4 September 2000	Priority Date 2 September 1999	
Applicant NEW ZEALAND PASTORAL AGRICULTURE RESEARCH INSTITUTE LIMITED et al			

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.
2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.
3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translations to those Offices.
4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices)(Article 39(1))(see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide

Name and mailing address of the IPEA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustalia.gov.au Facsimile No. (02) 6285 3929	Authorized officer GARETH COOK Telephone No. (02) 6283 2541
---	---

INTERNATIONAL COOPERATION TREATY
PCT
INTERNATIONAL PRELIMINARY EXAMINATION REPORT
(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 18134/8X109SB	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416).
International Application No. PCT/NZ00/00174	International Filing Date (<i>day/month/year</i>) 4 September 2000	Priority Date (<i>day/month/year</i>) 2 September 1999
International Patent Classification (IPC) or national classification and IPC Int. Cl. ⁷ C12N 15/31, A01H 5/00, A01N 63/02, C07K 14/24, C12N 15/70, C12N 15/82, C12Q 1/68		
Applicant NEW ZEALAND PASTORAL AGRICULTURE RESEARCH INSTITUTE LIMITED et al		

1.	This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2.	This REPORT consists of a total of 3 sheets, including this cover sheet. <input checked="" type="checkbox"/> This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT). These annexes consist of a total of 5 sheet(s).
3. This report contains indications relating to the following items:	
I	<input checked="" type="checkbox"/> Basis of the report
II	<input type="checkbox"/> Priority
III	<input type="checkbox"/> Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
IV	<input type="checkbox"/> Lack of unity of invention
V	<input checked="" type="checkbox"/> Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
VI	<input type="checkbox"/> Certain documents cited
VII	<input type="checkbox"/> Certain defects in the international application
VIII	<input type="checkbox"/> Certain observations on the international application

Date of submission of the demand 26 March 2001	Date of completion of the report 28 November 2001
Name and mailing address of the IPEA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaaustralia.gov.au Facsimile No. (02) 6285 3929	Authorized Officer GARETH COOK Telephone No. (02) 6283 2541

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/NZ00/00174

I. Basis of the report

1. With regard to the elements of the international application:*
- ☐ the international application as originally filed.
- ☒ the description, pages 1-2, 5-48, as originally filed,
pages , filed with the demand,
pages 3, 4, received on 5 October 2001 with the letter of 5 October 2001
- ☒ the claims, pages 52-54, as originally filed,
pages , as amended (together with any statement) under Article 19,
pages , filed with the demand,
pages 49-51, received on 23 November 2001 with the letter of 23 November 2001
- ☒ the drawings, pages 1-8, as originally filed,
pages , filed with the demand,
pages , received on with the letter of
- ☒ the sequence listing part of the description:
pages 1-46, as originally filed
pages , filed with the demand
pages , received on with the letter of
2. With regard to the language, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.
These elements were available or furnished to this Authority in the following language which is:
- ☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).
3. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:
- ☒ contained in the international application in written form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☒ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☒ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished
4. ☐ The amendments have resulted in the cancellation of:
- ☐ the description, pages
- ☐ the claims, Nos.
- ☐ the drawings, sheets/fig.
5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).

** Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/NZ00/00174

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1.	Statement		
	Novelty (N)	Claims 1-41	YES
		Claims	NO
	Inventive step (IS)	Claims 1-41	YES
		Claims	NO
	Industrial applicability (IA)	Claims 1-41	YES
		Claims	NO

2. Citations and explanations (Rule 70.7)

Novelty (N) and Inventive Step (IS) claims 1 to 41

The claims are to a an insecticidal protein complex, the polynucleotide encoding it and uses associated with protein and polypeptide. The closest prior art is considered to be Grkovic S *et al*, *Applied and Environmental Microbiology*, 1995, 61(6):2218-2223 which discloses the plasmid which encodes for the insecticidal protein complex. The document however does not disclose the sequence of the polypeptides or their encoding polynucleotides, hence the claims are considered novel. It is also considered that it would require more than routine effort to identify the genes encoding the protein complex, hence the claims are considered to involve an inventive step. As such the claims meet the requirements of Articles 33(2) and 33(3) of the PCT.

Industrial Applicability (IA) claims 1 to 41

Claims 1 to 41 are considered to be Industrially Applicable under Article 33(4) of the PCT.

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

PCT

To:

SIMS, ALLEN, FINCH, LEWIS, MURPHY,
ROGERS, TYRER-HARDING, WELLS,
29 Clarence Street
Private Bag 3140
Hamilton 2001, New Zealand
NEW ZEALAND

NOTIFICATION OF TRANSMITTAL OF
THE INTERNATIONAL SEARCH REPORT
OR THE DECLARATION

(PCT Rule 44.1)

Date of mailing
(day/month/year)

06/06/2001

Applicant's or agent's file reference

18134/8X109

FOR FURTHER ACTION

See paragraphs 1 and 4 below

International application No.

PCT/NZ 00/00174

International filing date
(day/month/year)

04/09/2000

Applicant

NEW ZEALAND PASTORAL AGRICULTURE RESEARCH INSTI...

1. ☒ The applicant is hereby notified that the International Search Report has been established and is transmitted herewith.

Filing of amendments and statement under Article 19:

The applicant is entitled, if he so wishes, to amend the claims of the International Application (see Rule 46):

When? The time limit for filing such amendments is normally 2 months from the date of transmittal of the International Search Report; however, for more details, see the notes on the accompanying sheet.

Where? Directly to the International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland
Facsimile No.: (41-22) 740.14.35

For more detailed instructions, see the notes on the accompanying sheet.

2. ☐ The applicant is hereby notified that no International Search Report will be established and that the declaration under Article 17(2)(a) to that effect is transmitted herewith.

3. ☐ With regard to the protest against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.

☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Further action(s):** The applicant is reminded of the following:

Shortly after **18 months** from the priority date, the International application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in Rules 90bis.1 and 90bis.3, respectively, before the completion of the technical preparations for international publication.

Within **19 months** from the priority date, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase until 30 months from the priority date (in some Offices even later).

Within **20 months** from the priority date, the applicant must perform the prescribed acts for entry into the national phase before all designated Offices which have not been elected in the demand or in a later election within 19 months from the priority date or could not be elected because they are not bound by Chapter II.

Name and mailing address of the International Searching Authority



European Patent Office, P.B. 5818 Patentlaan 2
NL-2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Mireille Claudepierre

NOTE FORM PCT/ISA/220

These Notes are intended to give the basic instructions concerning the filing of amendments under article 19. The Notes are based on the requirements of the Patent Cooperation Treaty, the Regulations and the Administrative Instructions under that Treaty. In case of discrepancy between these Notes and those requirements, the latter are applicable. For more detailed information, see also the PCT Applicant's Guide, a publication of WIPO.

In these Notes, "Article", "Rule", and "Section" refer to the provisions of the PCT, the PCT Regulations and the PCT Administrative Instructions respectively.

INSTRUCTIONS CONCERNING AMENDMENTS UNDER ARTICLE 19

The applicant has, after having received the international search report, one opportunity to amend the claims of the international application. It should however be emphasized that, since all parts of the international application (claims, description and drawings) may be amended during the international preliminary examination procedure, there is usually no need to file amendments of the claims under Article 19 except where, e.g. the applicant wants the latter to be published for the purposes of provisional protection or has another reason for amending the claims before international publication. Furthermore, it should be emphasized that provisional protection is available in some States only.

What parts of the international application may be amended?

Under Article 19, only the claims may be amended.

During the international phase, the claims may also be amended (or further amended) under Article 34 before the International Preliminary Examining Authority. The description and drawings may only be amended under Article 34 before the International Examining Authority.

Upon entry into the national phase, all parts of the international application may be amended under Article 28 or, where applicable, Article 41.

When?

Within 2 months from the date of transmittal of the international search report or 16 months from the priority date, whichever time limit expires later. It should be noted, however, that the amendments will be considered as having been received on time if they are received by the International Bureau after the expiration of the applicable time limit but before the completion of the technical preparations for international publication (Rule 46.1).

Where not to file the amendments?

The amendments may only be filed with the International Bureau and not with the receiving Office or the International Searching Authority (Rule 46.2).

Where a demand for international preliminary examination has been/is filed, see below.

How?

Either by canceling one or more entire claims, by adding one or more new claims or by amending the text of one or more of the claims as filed.

A replacement sheet must be submitted for each sheet of the claims which, on account of an amendment or amendments, differs from the sheet originally filed.

All the claims appearing on a replacement sheet must be numbered in Arabic numerals. Where a claim is cancelled, no renumbering of the other claims is required. In all cases where claims are renumbered, they must be renumbered consecutively (Administrative Instructions, Section 205(b)).

The amendments must be made in the language in which the international application is to be published.

What documents must/may accompany the amendments?

Letter (Section 205(b)):

The amendments must be submitted with a letter.

The letter will not be published with the international application and the amended claims. It should not be confused with the "Statement under Article 19(1)" (see below, under "Statement under Article 19(1)").

The letter must be in English or French, at the choice of the applicant. However, if the language of the international application is English, the letter must be in English; if the language of the international application is French, the letter must be in French.

NOTES TO FORM PCT/ISA/220 (continued)

The letter must indicate the differences between the claims as filed and the claims as amended. It must, in particular, indicate, in connection with each claim appearing in the international application (it being understood that identical indications concerning several claims may be grouped), whether

- (i) the claim is unchanged;
- (ii) the claim is cancelled;
- (iii) the claim is new;
- (iv) the claim replaces one or more claims as filed;
- (v) the claim is the result of the division of a claim as filed.

The following examples illustrate the manner in which amendments must be explained in the accompanying letter:

1. [Where originally there were 48 claims and after amendment of some claims there are 51]:
"Claims 1 to 29, 31, 32, 34, 35, 37 to 48 replaced by amended claims bearing the same numbers; claims 30, 33 and 36 unchanged; new claims 49 to 51 added."
2. [Where originally there were 15 claims and after amendment of all claims there are 11]:
"Claims 1 to 15 replaced by amended claims 1 to 11."
3. [Where originally there were 14 claims and the amendments consist in cancelling some claims and in adding new claims]:
"Claims 1 to 6 and 14 unchanged; claims 7 to 13 cancelled; new claims 15, 16 and 17 added." or
"Claims 7 to 13 cancelled; new claims 15, 16 and 17 added; all other claims unchanged."
4. [Where various kinds of amendments are made]:
"Claims 1-10 unchanged; claims 11 to 13, 18 and 19 cancelled; claims 14, 15 and 16 replaced by amended claim 14; claim 17 subdivided into amended claims 15, 16 and 17; new claims 20 and 21 added."

"Statement under article 19(1)" (Rule 46.4)

The amendments may be accompanied by a statement explaining the amendments and indicating any impact that such amendments might have on the description and the drawings (which cannot be amended under Article 19(1)).

The statement will be published with the international application and the amended claims.

It must be in the language in which the international application is to be published.

It must be brief, not exceeding 500 words if in English or if translated into English.

It should not be confused with and does not replace the letter indicating the differences between the claims as filed and as amended. It must be filed on a separate sheet and must be identified as such by a heading, preferably by using the words "Statement under Article 19(1)."

It may not contain any disparaging comments on the international search report or the relevance of citations contained in that report. Reference to citations, relevant to a given claim, contained in the international search report may be made only in connection with an amendment of that claim.

Consequence if a demand for international preliminary examination has already been filed

If, at the time of filing any amendments under Article 19, a demand for international preliminary examination has already been submitted, the applicant must preferably, at the same time of filing the amendments with the International Bureau, also file a copy of such amendments with the International Preliminary Examining Authority (see Rule 62.2(a), first sentence).

Consequence with regard to translation of the international application for entry into the national phase

The applicant's attention is drawn to the fact that, where upon entry into the national phase, a translation of the claims as amended under Article 19 may have to be furnished to the designated/elected Office, instead of, or in addition to, the translation of the claims as filed.

For further details on the requirements of each designated/elected Office, see Volume II of the PCT Applicant's Guide.

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 18134/8X109	FOR FURTHER ACTION <small>see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.</small>	
International application No. PCT/NZ 00/ 00174	International filing date (day/month/year) 04/09/2000	(Earliest) Priority Date (day/month/year) 02/09/1999
Applicant NEW ZEALAND PASTORAL AGRICULTURE RESEARCH INSTI...		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 6 sheets.
☒ It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**
 - a. With regard to the language, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

☐ the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).
 - b. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of the sequence listing :

☒ contained in the international application in written form.
☐ filed together with the international application in computer readable form.
☐ furnished subsequently to this Authority in written form.
☒ furnished subsequently to this Authority in computer readable form.
☒ the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
☒ the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished
2. ☒ **Certain claims were found unsearchable (See Box I).**
3. ☐ **Unity of invention is lacking (see Box II).**
4. With regard to the title,

☐ the text is approved as submitted by the applicant.
☒ the text has been established by this Authority to read as follows:

NUCLEOTIDES SEQUENCES ENCODING AN INSECTIDAL PROTEIN COMPLEX FROM SERRATIA
5. With regard to the abstract,

☒ the text is approved as submitted by the applicant.
☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.
6. The figure of the drawings to be published with the abstract is Figure No.

☐ as suggested by the applicant.
☐ because the applicant failed to suggest a figure.
☐ because this figure better characterizes the invention.

☒ **None of the figures.**

INTERNATIONAL SEARCH REPORT

International Application No

PC 00/00174

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/31 C12N15/70 C12N15/82 C07K14/24 C12Q1/68
A01N63/02 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K A01H C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

BIOSIS, EPO-Internal, STRAND, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>JACKSON T A ET AL: "PATHOGEN TO PRODUCT DEVELOPMENT OF SERRATIA-ENTOMOPHILA ENTEROBACTERIACEAE AS A COMMERCIAL BIOLOGICAL CONTROL AGENT FOR NEW ZEALAND GRASS GRUB COSTELYTRA-ZEALANDICA" JACKSON, T. A. AND T. R. GLARE (ED.). USE OF PATHOGENS IN SCARAB PEST, 1992, pages 191-198, XP000997900 0-946707-35-9. 1992 the whole document</p> <p style="text-align: center;">--- -/-</p>	32



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

Date of the actual completion of the international search

23 May 2001

Date of mailing of the international search report

06/06/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Holtorf, S

INTERNATIONAL SEARCH REPORT

International Application No

PO 00/00174

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GRKOVIC STEVE ET AL: "Genes Essential for Amber Disease in Grass Grubs Are Located on the Large Plasmid Found in Serratia entomophila and Serratia proteamaculans." APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 61, no. 6, 1995, pages 2218-2223, XP000994573 ISSN: 0099-2240 cited in the application the whole document ---	
A	GLARE TRAVIS R ET AL: "Plasmid transfer among several members of the family Enterobacteriaceae increases the number of species capable of causing experimental amber disease in grass grub." FEMS MICROBIOLOGY LETTERS, vol. 139, no. 2-3, 1996, pages 117-120, XP000998482 ISSN: 0378-1097 cited in the application the whole document ---	
A	WO 99 42589 A (NOVARTIS ERFIND VERWALT GMBH ;NOVARTIS AG (CH); KRAMER VANCE CARY) 26 August 1999 (1999-08-26) the whole document ---	
A	WO 98 08932 A (DOW AGROSCIENCES LLC ;WISCONSIN ALUMNI RES FOUND (US)) 5 March 1998 (1998-03-05) the whole document ---	
A	WO 98 08388 A (MORGAN JAMES ALUN WYNNE ;JARRETT PAUL (GB); ELLIS DEBORAH JUNE (GB) 5 March 1998 (1998-03-05) the whole document ---	
A	WO 97 17432 A (WISCONSIN ALUMNI RES FOUND) 15 May 1997 (1997-05-15) the whole document ---	
A	BOWEN D ET AL: "INSECTICIDAL TOXINS FROM THE BACTERIUM Photorhabdus limnescens" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, US, vol. 280, 26 June 1998 (1998-06-26), pages 2129-2132, XP002115650 ISSN: 0036-8075 cited in the application ---	
	--- -/--	

INTERNATIONAL SEARCH REPORT

International Application No

PC 00/00174

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NUNEZ-VALDEZ M E ET AL: "The amb2 locus from <i>Serratia entomophila</i> confers anti-feeding effect on larvae of <i>Costelytra zealandica</i> (Coleoptera: Scarabaeidae)" GENE, NL, ELSEVIER BIOMEDICAL PRESS. AMSTERDAM, vol. 172, no. 1, 12 June 1996 (1996-06-12), pages 75-79, XP004042712 ISSN: 0378-1119 cited in the application	
P, X	HURST MARK R H ET AL: "Plasmid-located pathogenicity determinants of <i>Serratia entomophila</i> , the causal agent of amber disease of grass grub, show similarity to the insecticidal toxins of <i>Photobacterium luminescens</i> ." JOURNAL OF BACTERIOLOGY, vol. 182, no. 18, September 2000 (2000-09), pages 5127-5138, XP002166799 ISSN: 0021-9193 the whole document	1-4, 9-16, 21-27, 31, 41

International Application No. PCT/NZ 00 00174

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 17

Present claim 17 relates to a ligand defined by reference to a desirable characteristic or property, namely binding to the polypeptide of claim 15:

The claim covers all products having this characteristic or property, whereas the application provides no support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for such products.

In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product by reference to a result to be achieved.

Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/NZ 00/00174**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 17
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/AU 00/00174

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9942589 A	26-08-1999	AU 3028699 A EP 1054972 A	06-09-1999 29-11-2000
WO 9808932 A	05-03-1998	AU 729228 B AU 1050997 A AU 2829997 A BR 9606889 A BR 9711441 A CA 2209659 A EP 0797659 A EP 0970185 A HU 9900768 A JP 2000515024 T PL 321212 A PL 332033 A SK 24699 A SK 93197 A TR 9901126 T WO 9717432 A	25-01-2001 29-05-1997 19-03-1998 28-10-1997 24-10-2000 15-05-1997 01-10-1997 12-01-2000 28-06-1999 14-11-2000 24-11-1997 16-08-1999 10-04-2000 06-05-1998 21-07-1999 15-05-1997
WO 9808388 A	05-03-1998	AU 4024997 A BR 9711285 A CN 1233938 A EP 0923295 A TR 9900435 T ZA 9707373 A	19-03-1998 17-08-1999 03-11-1999 23-06-1999 21-06-1999 15-02-1999
WO 9717432 A	15-05-1997	AU 729228 B AU 1050997 A BR 9606889 A CA 2209659 A EP 0797659 A HU 9900768 A PL 321212 A PL 332033 A SK 93197 A AU 2829997 A BR 9711441 A EP 0970185 A JP 2000515024 T SK 24699 A TR 9901126 T WO 9808932 A	25-01-2001 29-05-1997 28-10-1997 15-05-1997 01-10-1997 28-06-1999 24-11-1997 16-08-1999 06-05-1998 19-03-1998 24-10-2000 12-01-2000 14-11-2000 10-04-2000 21-07-1999 05-03-1998

PATENT COOPERATION TREATY
PCT
INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

REC'D 07 SEP 2001

PCT

Applicant's or agent's file reference 18134/8X109SB	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416).
International Application No. PCT/NZ00/00174	International Filing Date (<i>day/month/year</i>) 4 September 2000	Priority Date (<i>day/month/year</i>) 2 September 1999
International Patent Classification (IPC) or national classification and IPC Int. Cl. ⁷ C12N 15/31, A01H 5/00, A01N 63/02, C07K 14/24, C12N 15/70, C12N 15/82, C12Q 1/68		
Applicant NEW ZEALAND PASTORAL AGRICULTURE RESEARCH INSTITUTELIMITED et al		

1.	This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.																
2.	This REPORT consists of a total of 3 sheets, including this cover sheet. <input checked="" type="checkbox"/> This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT). These annexes consist of a total of 5 sheet(s).																
3.	This report contains indications relating to the following items: <table style="width: 100%; margin-top: 10px;"> <tr> <td style="width: 5%;">I</td> <td><input checked="" type="checkbox"/> Basis of the report</td> </tr> <tr> <td>II</td> <td><input type="checkbox"/> Priority</td> </tr> <tr> <td>III</td> <td><input type="checkbox"/> Non-establishment of opinion with regard to novelty, inventive step and industrial applicability</td> </tr> <tr> <td>IV</td> <td><input type="checkbox"/> Lack of unity of invention</td> </tr> <tr> <td>V</td> <td><input checked="" type="checkbox"/> Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement</td> </tr> <tr> <td>VI</td> <td><input type="checkbox"/> Certain documents cited</td> </tr> <tr> <td>VII</td> <td><input type="checkbox"/> Certain defects in the international application</td> </tr> <tr> <td>VIII</td> <td><input type="checkbox"/> Certain observations on the international application</td> </tr> </table>	I	<input checked="" type="checkbox"/> Basis of the report	II	<input type="checkbox"/> Priority	III	<input type="checkbox"/> Non-establishment of opinion with regard to novelty, inventive step and industrial applicability	IV	<input type="checkbox"/> Lack of unity of invention	V	<input checked="" type="checkbox"/> Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement	VI	<input type="checkbox"/> Certain documents cited	VII	<input type="checkbox"/> Certain defects in the international application	VIII	<input type="checkbox"/> Certain observations on the international application
I	<input checked="" type="checkbox"/> Basis of the report																
II	<input type="checkbox"/> Priority																
III	<input type="checkbox"/> Non-establishment of opinion with regard to novelty, inventive step and industrial applicability																
IV	<input type="checkbox"/> Lack of unity of invention																
V	<input checked="" type="checkbox"/> Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement																
VI	<input type="checkbox"/> Certain documents cited																
VII	<input type="checkbox"/> Certain defects in the international application																
VIII	<input type="checkbox"/> Certain observations on the international application																

Date of submission of the demand 26 March 2001	Date of completion of the report 28 November 2001
Name and mailing address of the IPEA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustalia.gov.au Facsimile No. (02) 6285 3929	Authorized Officer GARETH COOK Telephone No. (02) 6283 2541

I. Basis of the report

1. With regard to the **elements** of the international application:*
- ☐ the international application as originally filed.
- ☒ the description, pages **1-2, 5-48**, as originally filed,
pages , filed with the demand,
pages **3, 4**, received on **5 October 2001** with the letter of **5 October 2001**
- ☒ the claims, pages **52-54**, as originally filed,
pages , as amended (together with any statement) under Article 19,
pages , filed with the demand,
pages **49-51**, received on **23 November 2001** with the letter of **23 November 2001**
- ☒ the drawings, pages **1-8**, as originally filed,
pages , filed with the demand,
pages , received on with the letter of
- ☒ the sequence listing part of the description:
pages **1-46**, as originally filed
pages , filed with the demand
pages , received on with the letter of
2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.
These elements were available or furnished to this Authority in the following language which is:
- ☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).
3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:
- ☒ contained in the international application in written form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☒ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☒ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished
4. ☐ The amendments have resulted in the cancellation of:
- ☐ the description, pages
- ☐ the claims, Nos.
- ☐ the drawings, sheets/fig.
5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).

** Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Claims 1-41	YES
	Claims	NO
Inventive step (IS)	Claims 1-41	YES
	Claims	NO
Industrial applicability (IA)	Claims 1-41	YES
	Claims	NO

2. Citations and explanations (Rule 70.7)

Novelty (N) and Inventive Step (IS) claims 1 to 41

The claims are to a an insecticidal protein complex, the polynucleotide encoding it and uses associated with protein and polypeptide. The closest prior art is considered to be Grkovic S *et al*, *Applied and Environmental Microbiology*, 1995, 61(6):2218-223 which discloses the plasmid which encodes for the insecticidal protein complex. The document however does not disclose the sequence of the polypeptides or their encoding polynucleotides, hence the claims are considered novel. It is also considered that it would require more than routine effort to identify the genes encoding the protein complex, hence the claims are considered to involve an inventive step. As such the claims meet the requirements of Articles 33(2) and 33(3) of the PCT.

Industrial Applicability (IA) claims 1 to 41

Claims 1 to 41 are considered to be Industrially Applicable under Article 33(4) of the PCT.

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
8 March 2001 (08.03.2001)

PCT

(10) International Publication Number
WO 01/16305 A2

- (51) International Patent Classification⁷: C12N 15/00 (74) Agents: SIMS, Anthony, W. et al.; 29 Clarence Street, Private Bag 3140, Hamilton 2001 (NZ).
- (21) International Application Number: PCT/NZ00/00174
- (22) International Filing Date:
4 September 2000 (04.09.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
337610 2 September 1999 (02.09.1999) NZ
- (71) Applicant (for all designated States except US): NEW ZEALAND PASTORAL AGRICULTURE RESEARCH INSTITUTE LIMITED [NZ/NZ]; 5th floor, Tower Block, Ruakura Research Centre, East Street, Hamilton 2001 (NZ).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): GLARE, Travis, Robert [AU/NZ]; 38 Whincorps Road, Halswell, Christchurch 8003 (NZ). HURST, Mark, Robin, Holmes [NZ/NZ]; 148 Hendersons Road, Hoon Hay, Christchurch 8002 (NZ). JACKSON, Trevor, Anthony [NZ/NZ]; 407 Halswell Road, Halswell, Christchurch 8003 (NZ).
- Published:
— Without international search report and to be republished upon receipt of that report.
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/16305 A2

(54) Title: NUCLEOTIDE SEQUENCES

(57) Abstract: The present invention concerns novel nucleotide sequences encoding proteins from the Enterobacteriaceae, *Serratia entomophila* and *Serratia proteamaculans*, and the use of said nucleotide sequences and proteins for inherent insecticidal and potentially metazoacidal properties. The invention relates to an isolated nucleic acid molecule comprising a nucleotide sequence that encodes an insecticidal protein complex, or a functional fragment, neutral mutation, or homolog thereof capable of hybridising with the nucleic acid molecule under standard hybridisation conditions. The nucleotide sequences include a pathogenicity-encoding region cloned from bacteria *Serratia entomophila* and *S. proteamaculans*. The region contain pathogenic determinants of a disease that affect the grass grub, *Costelytra zealandica* Coleoptera: Scarabaeidae, an important insect pasture pest in New Zealand. The proteins encoded by determined genes may be used for insect control whether as an inundative pesticide, within baits or expressed in other organisms such as plants or microbes.

NUCLEOTIDE SEQUENCES

TECHNICAL FIELD

8/p1b

The present invention concerns novel nucleotide sequences encoding insecticidal proteins from the Enterobacteriaceae, *Serratia entomophila* and *Serratia proteamaculans*, and the use of said nucleotide sequences and insecticidal proteins.

BACKGROUND ART

Some *Serratia entomophila* and *Serratia proteamaculans* strains in New Zealand are known to cause a disease in the major scarab pest, *Costelytra zealandica* (New Zealand grass grub). The disease was first discovered and described by Trought and Jackson (1982) and was later named amber disease after the distinctive colour of affected insects (Stucki et al. 1984). One species capable of causing the disease, *Serratia entomophila*, was developed into a commercially-available product ("Invade") in 1989.

The disease is highly host specific, only known to infect a single indigenous species of New Zealand scarab larva. The disease appears unique among insects and results not from rapid invasion of the haemocoel, but from a slow colonisation of the gut. The disease has a distinct phenotypic progression, with infected hosts ceasing feeding within 2-5 days of ingesting pathogenic cells. The normally black gut clears around this time (Jackson et al. 1993) and the levels of the major gut digestive enzymes (trypsin and so forth) decrease sharply (Jackson, 1995). The clearance of the gut results in a characteristic amber colour of the infected hosts. The larvae may remain in this state for a prolonged period (1-3 months) before bacteria eventually invade the haemocoel, causing rapid death.

The finding of a plasmid that apparently encoded the disease was reported in Glare et al. (1993) by showing a correlation between pADAP presence and disease occurrence in

bacterial strains. This was further confirmed by Glare et al. (1996) who showed that transfer of the plasmid from pathogenic to non-pathogenic strains resulted in a change to pathogenic.

5 Grkovic et al. (1995) showed that disruption of the plasmid by transposon insertion could alter pathogenicity without fully defining the area containing the gene cassette. By marker exchange, they showed that a 10.5kb *HindIII* (pGLA20) construct from pADAP encoded some functions of amber disease. However, the clone did not contain all disease encoding plasmid-borne regions.

10 Another region involved in amber disease encoding was located by Nunez-Valdez and Mahanty (1996). They located a locus, *amb2*, by transposon mutagenesis and searching a cosmid genomic library. This region was chromosomally located and was involved in antifeeding in the larvae of *Costelytra zealandica*. However, the current applicant's research has demonstrated that the *amb2* region is located on pADAP remote from the virulence gene and is probably regulatory in function.

15 Insecticidal toxins which share some protein homology to the *Serratia* insecticidal proteins of the present invention have been recently discovered (PCT/US96/18803; PCT/US97/07657) by a group at Wisconsin University (Blackburn et al. 1998; Bowen et al. 1998; Bowen and Ensign 1998). These insecticidal toxins are produced from a gene region in *Photorhabdus luminescens* which resembles the *Serratia* virulence region in the
20 clustering of the genes and at the protein level, but has very little DNA homology with the *Serratia* genes. They have shown high molecular weight proteins from *Photorhabdus luminescens* are insecticidal to a number of insects from different orders. The lack of DNA homology over the majority of the region, as opposed to protein homology, between the *Serratia* genes and *Photorhabdus* genes suggests that these proteins have evolved as a
25 result of convergent evolution leading to the formation of a distinct protein family with a

common function.

The present applicant has now found that three regions of the pADAP plasmid are required for full insecticidal function. Sequence analysis of these three regions has shown that the present applicant has isolated and identified a novel toxin from *Serratia* species that
5 belongs to a new family of insecticidal toxins. It is broadly to this toxin that the present invention is directed.

DISCLOSURE OF INVENTION

According to a first aspect of the present invention, there is provided an isolated nucleic acid molecule comprising a nucleotide sequence of SEQ ID NO: 1 which encodes an
10 insecticidal protein complex, or a functional fragment, neutral mutation, or homolog thereof which have at least 75% nucleic acid homology to SEQ ID NO: 1 and are capable of hybridising with said nucleic acid molecule under stringent hybridisation conditions.

The invention also provides an isolated nucleic acid molecule comprising the nucleotide sequence 1995-18937 of SEQ ID NO: 1 which encodes an insecticidal protein complex, or
15 a functional fragment, neutral mutation, or homolog thereof capable of hybridising with said nucleic acid molecule under standard hybridisation conditions.

The invention also provides an isolated nucleic acid molecule comprising one or more of the nucleotide sequences 2411-9547, 9589-13883 or 14546-17467 of SEQ ID NO: 1 which
20 encode insecticidal proteins, or a functional fragment, neutral mutation, or homolog thereof capable of hybridising with said nucleic acid molecule under standard hybridisation conditions.

Preferably the nucleic acid molecule comprises all of nucleotide sequences 2411-9547, 9598-13884 and 14546-17467 of SEQ ID NO: 1.

The invention further relates to an isolated nucleic acid molecule comprising a sequence of SEQ ID NO: 1, nucleotides 1955-18937 of SEQ ID NO: 1 or one or more of nucleotides 2411-9547, 9598-13884 or 14546-17467 of SEQ ID NO: 1, operably linked to at least one further nucleotide sequence which encode an insecticidal protein. For example, the at least
5 one further nucleotide sequence may be the nucleotide sequence which codes for the *Bacillus delta* endo toxins, vegetative insecticidal proteins (vips), cholesterol oxidases, *Clostridium bifermentens* mosquitocidal toxins and/or *Photorhabdus luminescens* toxins and so forth.

The nucleic acid molecule may comprise DNA, cDNA or RNA.

- 10 Preferably said fragment, neutral mutation or homolog thereof is capable of hybridising to said nucleic acid molecule under stringent hybridisation conditions.

The invention further relates to nucleic acid molecules which hybridise to the nucleotide sequence of SEQ ID NO: 1, or nucleotides 1955-18937, 2411-9547, 9598-13884 or 14546-17467 of SEQ ID NO: 1 if there is at least 75% or greater identity between the sequences.

- 15 The nucleic acid molecule may be isolated from *Serratia entomophila* or *Serratia proteamaculans* strains.

Also provided by the present invention are recombinant expression vectors containing the nucleic acid molecule of the invention and hosts transformed with the vector of the invention capable of expressing a polypeptide of the invention.

- 20 The vector may be selected from any suitable natural or artificial plasmid/vector. For example, pUC 19 (Yannish-Perron et al. 1995), pProEX HT (GibcoBRL, Gaithersburg, MD, USA), pBR322 (Bolivar et al. 1977), pACYC184 (Chang et al. 1978), pLAFR3 (Staskowicz et al. 1987), and so forth.

In a further aspect, the invention provides a method of producing a polypeptide of the invention comprising the steps of:

- (a) culturing a host cell which has been transformed or transfected with a vector as defined above to express the encoded polypeptide or peptide; and
- 5 (b) recovering the expressed polypeptide or peptide.

An additional aspect of the present invention provides a ligand that binds to a polypeptide of the invention. Most usually, the ligand is an antibody or antibody binding fragment. Such ligands also form a part of this invention.

According to a further aspect of the present invention there are provided probes and primers
10 comprising a fragment of the nucleic acid molecule of the invention capable of hybridising under stringent conditions to a native insecticidal gene sequence. Such probes and primers are useful, for example, in studying the structure and function of this novel gene and for obtaining homologs of the gene from bacteria other than *Serratia* sp.

According to a still further aspect of the present invention there is provided a polypeptide
15 having insecticidal activity encoded by the nucleic acid molecule of the invention, or a functional fragment, neutral mutation or homolog thereof.

The polypeptide may comprise the amino acid sequence of SEQ ID NO: 1 or a functional fragment, neutral mutation or homolog thereof.

The polypeptide may comprise amino acids 32-5118 of SEQ ID NO: 1.

20 The polypeptide may comprise at least one amino acid sequence of SEQ ID NO: 2; SEQ ID NO: 3; SEQ ID NO: 4; SEQ ID NO: 5 or SEQ ID NO: 6.

Preferably the polypeptide comprises amino acid sequence SEQ ID NO: 4; SEQ ID NO: 5

and SEQ ID NO: 6.

More preferably the polypeptide comprises all of SEQ ID NOs: 2-6.

Conveniently, the polypeptide of the invention is obtained by expression of a DNA sequence coding therefore in a host cell or organism.

- 5 The polypeptide may comprise the amino acid sequence of SEQ ID NO: 1 linked to at least one further amino acid sequence encoding an insecticidal protein. For example, the at least one further amino acid sequence may be the amino acid sequence which codes for *Bacillus* delta endo toxins, vegetative insecticidal proteins (vips), cholesterol oxidases, *Clostridium bifermentens* mosquitocidal toxins and/or *Photorhabdus luminescents* toxins etc.
- 10 The invention further relates to polypeptides comprising at least 50%, preferably 60%, more preferably 70% and most preferably 90-95% or greater identity to SEQ ID NO: 1.
- The polypeptide may be produced by expression of a vector comprising the nucleic acid molecule of the invention or a functional fragment, neutral mutation or homolog thereof, in a suitable host cell.
- 15 According to a further aspect, there is provided an insecticidal composition comprising at least the polypeptide of the invention and an agriculturally acceptable carrier such as would be known to a person skilled in the art. More than one polypeptide of the invention can of course, be included in the composition. In addition, the composition may comprise one or more additional pesticides, for example, compounds known to possess herbicidal,
- 20 fungicidal, insecticidal or nematocidal activity.

The composition may further comprise other known insecticidally active agents, such as *Bacillus* delta endo toxins, vegetative insecticidal proteins (vips), cholesterol oxidases, *Clostridium bifermentens* mosquitocidal toxins and/or *Photorhabdus luminescents* toxins

and so forth.

According to a further aspect, there is provided a method of combating pests, especially insects at a locus or host for the pest infested with or liable to be infested therewith, said method comprising applying to a locus, host and/or the pest, an effective amount of the polypeptide of the invention that has functional insecticidal activity against said pest.

According to a further aspect the invention provides a method of inducing amber disease or like condition in insects comprising delivery to an insect an effective amount of the polypeptide of the invention that has functional insecticidal activity against said insect.

The insect may be selected from the order comprising Coleoptera (such as the black beetle, *Heteronychus arator* (F.), or the black vine weevil, *Otiorhynchus sulcatus* (F.)); Dictyoptera (eg. The German cockroach, *Blattella germanica* (L.), or the subterranean termite *Coptotermes* spp.); Diptera (eg. the housefly *Musca domestica* L. or the blowfly *Lucilia cuprina* (Wiedermann); Orthoptera (eg. The black field cricket *Telleogryllus commodus* (Walker) or the migratory locust *Locusta migratoria* L.); Hymenoptera (eg. The German wasp, *Vespula germanica* F.); Hemiptera (such as the green vegetable bug *Nezara viridula* (L.) or the green peach aphid *Myzus persicae* (Sulzer)) the Lepidoptera (eg. the tomato fruitworm, *Helicoverpa armigera* (Walker), or the codling moth, *Laspeyresia pomonella* (L.)).

The insecticidal polypeptide may be delivered to the insect orally either as a solid bait matrix, as a sprayable insecticide sprayed onto a substrate upon which the insect feeds, applied directly to the soil subsurface or as a drench or is expressed in an transgenic plant, bacterium, virus or fungus upon which the insect feeds, or by any other suitable method which would be obvious to a person skilled in the art.

According to a further aspect, the invention provides a transgenic plant, bacterium virus or

fungus, incorporating in its genome, a nucleic acid molecule of the invention providing the plant, bacterium virus or fungus with an ability to express an effective amount of an insecticidal polypeptide.

DEFINITIONS AND METHODS

- 5 The following definitions and methods are provided to better define the present invention and to guide those of ordinary skill in the art in the practice of the present invention.

Definitions of common terms in molecular biology may also be found in Lewin, *Genes V*, Oxford University Press: New York, 1994.

- 10 The term "native" refers to a naturally-occurring nucleic acid or polypeptide, including, wild-type sequence and alleles thereof.

A "homolog" has at least one of the biological activities of the nucleic acid or polypeptide of the invention and comprises at least 50-70% identical amino acid or nucleic acid sequence thereto, preferably 75-85% and most preferably 90-95% identical amino acid or nucleic acid sequence thereto.

- 15 The term "neutral mutation" means a mutation, (that is - a change in the nucleotide or polypeptide sequence such as by deletion, substitution, inversion or insertion, any of which have no effect on the function of the encoded protein).

- 20 As indicated above, also possible are variants of the polypeptide or peptide that differ from the native amino acid sequence by insertion, substitution or deletion of one or more amino acids. Where such a variant is desired, the nucleotide sequence of the native DNA is altered appropriately. This alteration can be made through elective synthesis of the DNA, or by modification of the native DNA by, for example, site specific or cassette mutagenesis. Preferably, where portions of cDNA or genomic DNA require sequence modifications, site-

specific primer directed mutagenesis is employed using techniques standard in the art.

In a further aspect, the present invention consists in replicable transfer vector suitable for use in preparing a polypeptide of the invention. These vectors may be constructed according to techniques well known in the art, or may be selected from cloning vecotrs
5 available in the art.

The cloning vector may be selected according to the host or host cell to be used. Useful vectors will generally have the following characteristics:

- (a) the ability to self-replicate;
- (b) the possession of a single target for any particular restriction endonuclease; and
- 10 (c) desirably, carry genes for a readily selectable marker such as antibiotic resistance.

Two major types of vector possessing these characteristics are plasmids and bacterial viruses (bacteriophages or phages). Presently preferred vectors include plasmids pMOS-Blue, pGem-T and pUC8.

The nucleic acids of the present invention can be free in solution, or attached by
15 conventional means to a solid support, or present in an expression vector or any other type of plasmid.

The term "isolated" means substantially separated or purified away from contaminating sequences in the cell or organism in which the nucleic acid naturally occurs and includes nucleic acids purified by standard purification techniques as well as nucleic acids prepared
20 by recombinant technology and those chemically synthesised.

The terms "DNA construct" means a construct incorporating the nucleic acid molecule of the present invention, or a fractional fragment, neutral mutation or homolog thereof in a

position whereby the protein coding sequence is under the control of an operably linked promoter capable of expression in a plant cell. Such promoters are well known in the art.

A fragment of a nucleic acid molecule according to the present invention is a portion of the nucleic acid that is less than full length and comprises at least a minimum length capable of
5 hybridising specifically with a nucleic acid molecule according to the present invention (or a sequence complementary thereto) under stringent conditions as defined below. A fragment according to the present invention has at least one of the biological activities of the nucleic acid or polypeptide of the present invention.

Nucleic acid probes and primers can be prepared based on nucleic acids according to the
10 present invention (for example, the sequence of SEQ ID NO: 1). A "probe" comprises an isolated nucleic acid attached to a detectable label or reporter molecule well known in the art. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.

"Primers" are short nucleic acids, preferably DNA oligonucleotides 15 nucleotides or more
15 in length, which are annealed to a complementary target DNA strand by nucleic acid hybridisation to form a hybrid between the primer and the target DNA strand, then extended along the target DNA strand by a polymerase, preferably a DNA polymerase. Primer pairs can be used for amplification of a nucleic acid sequence, (for example, by the polymerase chain reaction (PCR) or other nucleic acid amplification methods well known
20 in the art). PCT-primer pairs can be derived from the sequence of a nucleic acid according to the present invention, (for example, by using computer programs intended for that purpose such as Primer (Version 0.5© 1991, Whitehead Institute for Biomedical Research, Cambridge, MA)).

Methods for preparing and using probes and primers are described, for example, in
25 Sambrook et al. *Molecular Cloning: A Laboratory Manual*, 2nd ed, vol. 1-3, ed Sambrook

et al. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY, 1989.

Probes or primers can be free in solution or covalently or noncovalently attached to a solid support by standard means.

The term "operably linked" means a first nucleic acid sequence linked to a second nucleic acid sequence when the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein coding regions, in reading frame.

10 The DNA molecules of the invention may be expressed by placing them in operable linkage with suitable control sequences in a replicable expression vector. Control sequences may include origins of replication, a promoter, enhancer and transcriptional terminator sequences, amongst others. The selection of the control sequence to be included in the expression vector is dependent on the type of host or host cell intended to be used for
15 expressing the DNA.

A "recombinant" nucleic acid is one that has a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic
20 acids (for example, by genetic engineering techniques).

Techniques for nucleic acid manipulation are described generally in, for example, Sambrook et al. (1989).

Large amounts of a nucleic acid according to the present invention can be produced by recombinant means well known in the art or by chemical synthesis.

Natural or synthetic nucleic acids according to the present invention can be incorporated into recombinant nucleic acid constructs, typically DNA constructs, capable of introduction into and replication in a host cell. Usually the DNA constructs will be suitable for replication in a unicellular host, such as *E. coli* or other commonly used bacteria, but can
5 also be introduced into yeast, mammalian, plant or other eukaryotic cells.

Preferably, such a nucleic acid construct is a vector comprising a replication system recognised by the host. For the practice of the present invention, well known compositions and techniques for preparing and using vectors, host cells, introduction of vectors into host cells and so forth., are employed, as discussed, *inter alia*, in Sambrook et al (1989).

10 A cell, tissue, organ, or organism into which has been introduced a foreign nucleic acid, such as a recombinant vector, is considered "transformed" or "transgenic". The DNA construct comprising a DNA sequence according to the present invention that is present in a transgenic host cell, particularly a transgenic plant, is referred to as a "transgene". The term "transgenic" or "transformed" when referring to a cell or organism, also includes;

- 15 (1) progeny of the cell or organism, and
- (2) plants produced from a breeding program employing such a "transgenic" plant as a parent in a cross and exhibiting an altered phenotype resulting from the presence of the recombinant DNA construct.

Generally, procaryotic, yeast, insect, or mammalian cells are useful hosts. Also included
20 within the term hosts are plasmid vectors. Suitable procaryotic hosts include *E. coli*, *Bacillus* species and various species of *Pseudomonas*. Commonly used promoters such as β -lactamase (penicillinase) and lactose (lac) promoter systems are all well known in the art. Any available promoter system compatible with the host of choice can be used. Vectors used in yeast are also available and well known. A suitable example is the 2 micron origin

of replication plasmid.

Similarly, vectors for use in mammalian cells are also well known. Such vectors include well known derivatives of SV-40, adenovirus, retrovirus-derived DNA sequences, *Herpes simplex* virus, and vectors derived from a combination of plasmid and phage DNA.

- 5 Further eucaryotic expression vectors are known in the art (for example in P.J. Southern and P. Berg, *J. Mol. Appl. Genet.* 1 327-341 (1982); S. Subramani et al., *Mol. Cell. Biol.* 1, 854-864 (1981); R.J. Kaufmann and P.A. Sharp, "Amplification and Expression of Sequences Cotransfected with a Modular Dihydrofolate Reducase Complementary DNA Gene, *J. Mol. Biol.* 159, 601-621 (1982); R.J. Kaufmann and P.A. Sharp, *Mol. Cell. Biol.* 10 159, 601-664 (1982); S.I. Scahill et al., "Expressions and Characterisation of the Product of a Human Immune Interferon DNA Gene in Chinese Hamster Ovary Cells," *Proc. Natl. Acad. Sci. USA.* 80, 4654-4659 (1983); G. Urlaub and L.A. Chasin, *Proc. Natl. Acad. Sci. USA.* 77, 4216-4220, (1980).

- The expression vectors useful in the present invention contain at least one expression control sequence that is operatively linked to the DNA sequence or fragment to be expressed. The control sequence is inserted in the vector in order to control and to regulate the expression of the cloned DNA sequence. Examples of useful expression control sequences are the lac system, the trp system, the tac system, the trc system, major operator and promoter regions of phage lambda, the glycolytic promoters of yeast acid phosphatase, 15 (for example, Pho5), the promoters of the yeast alpha-mating factors, and promoters derived from polyoma, adenovirus, retrovirus, and simian virus (for example, the early and late promoters of SV-40), and other sequences known to control the expression of genes of prokaryotic and eucaryotic cells and their viruses or combinations thereof.

- In the construction of a vector it is also an advantage to be able to distinguish the vector 25 incorporating the foreign DNA from unmodified vectors by a convenient and rapid assay.

Reporter systems useful in such assays include reported genes, and other detectable labels which produce measurable colour changes, antibiotic resistance and the like. In one preferred vector, the β -galactosidase reporter gene is used, which gene is detectable by clones exhibiting a blue phenotype on X-gal plates. This facilitates selection. In one
5 embodiment, the β -galactosidase gene may be replaced by a polyhedrin-encoding gene; which gene is detectable by clones exhibiting a white phenotype when stained with X-gal.

This blue-white colour selection can serve as a useful marker for detecting recombinant vectors.

Once selected, the vectors may be isolated from the culture using routine procedures such
10 as freeze-thaw extraction followed by purification.

For expression, vectors containing the DNA of the invention to be expressed and control signals are inserted or transformed into a host or host cell. Some useful expression host cells include well-known prokaryotic and eucaryotic cells. Some suitable prokaryotic hosts include, for example, *E. coli*, such as *E. coli*, S G-936, *E. coli* HB 101, *E. coli* W3110, *E.*
15 *coli* X1776, *E. coli*, X2282, *E. coli* DHT and *E. coli* MR01, *Pseudomonas*, *Bacillus*, such as *Bacillus subtilis* and *Streptomyces*. Suitable eucaryotic cells include yeast and other fungi, insect, animal cells, such as COS cells and CHO cells, human cells and plant cells in tissue culture.

Depending on the host used, transformation is performed according to standard techniques
20 appropriate to such cells. For prokaryotes or other cells that contain substantial cell walls, the calcium treatment process (Cohen, S N *Proceedings, National Academy of Science, USA* 69 2110 (1972)) may be employed. For mammalian cells without such cell walls the calcium phosphate precipitation method of Graeme and Van Der Eb, *Virology* 52:546 (1978) is preferred. Transformations into plants may be carried out using *Agrobacterium*
25 *tumefaciens* (Shaw et al., *Gene* 23:315 (1983)) or into yeast according to the method of Van

Solingen et al. *J. Bact.* 130:946 (1977) and Hsiao et al. *Proceedings, National Academy of Science*, 76:3829 (1979).

Upon transformation of the selected host with an appropriate vector the polypeptide, or peptide encoded can be produced, often in the form of fusion protein, by culturing the host cells. The polypeptide, or peptide, of the invention may be detected by rapid assays as indicated above. The polypeptide, or peptide, is then recovered and purified as necessary. Recovery and purification can be achieved using any of those procedures known in the art, for example by absorption onto the elution from an anion exchange resin. This method of producing a polypeptide, or peptide, of the invention constitutes a further aspect of the present invention.

Host cells transformed with the vectors of the invention also form a further aspect of the present invention.

Methods for chemical synthesis of nucleic acids are well known and can be performed, for example, on commercial automated oligonucleotide synthesisers.

The term "stringent conditions" is functionally defined with regard to the hybridisation of a nucleic acid probe to a target nucleic acid (for example, to a particular nucleic acid sequence of interest) by the hybridisation procedure discussed in Sambrook et al. (1989) at 9.52-9.55 and 9.56-9.58.

Regarding the amplification of a target nucleic acid sequence (for example, by PCR) using a particular amplification primer pair, stringent conditions are conditions that permit the primer pair to hybridise only to the target nucleic acid sequence to which a primer having the corresponding wild type sequence (or its complement) would bind.

Nucleic acid hybridisation is affected by such conditions as salt concentration, temperature, or organic solvents, in addition to the base composition, length of the complementary

strands, and the number of nucleotide base mismatches between the hybridising nucleic acids, as will be readily appreciated by those skilled in the art.

When referring to a probe or primer, the term "specific for (a target sequence)" indicates that the probe or primer hybridises under stringent conditions only to the target sequence in
5 a given sample comprising the target sequence.

The term "protein (or polypeptide)" refers to a protein encoded by the nucleic acid molecule of the invention including fragments, mutations and homologs having the same biological activity (for example, insecticidal activity). The polypeptide of the invention can be isolated from a natural source, produced by the expression of a recombinant nucleic acid
10 molecule or be chemically synthesised.

Peptides having substantial sequence identity to the above-mentioned peptides can also be employed in preferred embodiments. Here, "substantial sequence identity" means that two peptide sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default gap weights, share at least 80% sequence identity, preferably at least 90%
15 sequence identity, more preferably at least 95% sequence identity or more. Preferably, residue positions that are not identical differ by conservative amino acid substitutions. For example, the substitution of amino acids having similar chemical properties such as charge or polarity are not likely to effect the properties of a protein. Examples include glutamine for asparagine, or glutamic acid for aspartic acid.

20 **BRIEF DESCRIPTION OF DRAWINGS**

The invention will be further defined by reference to the specification and the following examples and figures herein.

Figure 1 shows restriction maps of clones used to isolate the pathogenic region and maps of the two pathogenic variants pMH32 and pMH41, in accordance

with a preferred embodiment of the present invention; and

Figure 2 shows deletion derivatives used in the study, restriction maps of the mutated constructs and recombinants, the phenotype of each mutation, the schematic diagram of the sequenced region, and a nucleotide sequence in accordance with a preferred embodiment of the present invention; and

Figure 3 shows hydrophobicity plots of SepC and its closest homologue TccC, in accordance with a preferred embodiment of the present invention; and

Figure 4 shows the comparison of protein sequences of the SepA and *P. luminescens* toxins, TcdA, TcaB and TccB Putative RGD motif is boxed, plus the site of proteolytic cleavage is illustrated, in accordance with a preferred embodiment of the present invention; and

Figure 5 shows the comparison of protein sequences of the SepC and *P. luminescens* toxin TccC, in accordance with a preferred embodiment of the present invention; and

Figure 6 shows the plasmid pADAP, in accordance with a preferred embodiment of the present invention.

BEST MODES FOR CARRYING OUT THE INVENTION

The invention will be further defined by reference to the specification and the following examples and figures herein in the ensuing description by way of example only where:


Figure 1 shows restriction maps of clones used to isolate the pathogenic region and maps of the two pathogenic variants pMH32 and pMH41, where:

(A) Is the pADAP *HindIII* clone pGLA-20 showing locations of the pGLA-20 mutations –

10, -13, and 35, which when recombined back into pADAP and bioassayed against grass grub, result in either a pathogenic phenotype, shown by full flag, or a healthy but non-feeding phenotype indicated by half filled flag. Map of pBG35 showing relative position of pGLA-20-35 mutation and the location of the 2.2kb *EcoRI* used as a probe to screen the

5 pADAP *BamHI* library; and

(B) Illustrated restriction enzyme maps of the pathogenic clones pMH32 and pMH41, area of deletion is indicated by Δ .

 pBR322 vector DNA;

 pLAFR3 vector DNA.

10 Restriction enzymes are abbreviated as follows: B, *BamHI*, Bg, *BglII*; E, *EcoRI*; H, *HindIII*; and X, *XbaI*.

Figure 2 shows:

(A) Which are Mini-Tn10 pACYC184 based deletion derivatives used in the study.

 is the pACYC184 vector,

15 Δ indicates deletion + pathogenic,

- loss of pathogenicity; and

(B) Illustrates restriction maps of the mutated constructs pBM32 and the pADK recombinants; and

(C) Where the phenotype of each mutant is indicated by flags.

20 Blocked flags indicates mutations that did not affect the disease process.


Open flags indicate mutations that abolish disease symptoms.

Half-filled flags denote mutations that abolish visual disease symptoms but are unable to feed.


* indicates pADK mutations obtained by Grkovic et al. (1995).

5 Restriction enzymes are abbreviated as follows: B, *Bam*HI, Bg, *Bgl*II; E, *Eco*RI; H, *Hind*III; and X, *Xba*I.

(D) Is a schematic diagram of the sequenced region, where:

 Denotes sequenced region.

Arrows indicate ORFs and their direction

 region homologous to spvB ... location of repeat.

10 (E) Is a nucleotide sequence of the 5 times 12bp repeat and the palindrome.

Restriction enzymes are abbreviated as follows: B, *Bam*HI, Bg, *Bgl*II; E, *Eco*RI; H, *Hind*III; and X, *Xba*I.

In Figure 3 hydrophobicity plots of SepC and its closest homologue TccC are shown. The scale is disproportional to size and has a scanning window of 17 amino-acid residues.

15 Figure 4 shows the comparison of protein sequences of the SepA and *P. luminescens* toxins, TcdA, TcaB and TccB. Putative RGD motif is boxed. The site of proteolytic cleavage is reported by Bowen et al. (1998) (Residue 1933 of TcdA) is indicated by an arrow.

Figure 5 shows the comparison of protein sequences of the SepC and *P. luminescens* toxin
20 TccC; and Figure 6 shows the plasmid pADAP.

PROTOCOL

Bacterial isolates and methods of culture

Table 1 lists bacterial isolates and plasmids used in the present invention. Bacteria were grown in LB broth or on LB agar (Sambrook et al. 1989), at 37° for *Escherichia coli* and
5 30°C for *S. entomophila*. Antibiotic concentrations used (µg/ml) for *Serratia* were kanamycin 100, chloramphenicol 90, tetracycline 30 and for *E. coli* strains were kanamycin 50, chloramphenicol 30, tetracycline 15, and ampicillin 100.

DNA isolation and manipulations

pADAP DNA was isolated from a 50ml overnight culture of bacteria using QIAGEN®
10 plasmid maxi kit (Qiagen, Hilden, Germany), as per the manufacturer's instructions. Standard DNA techniques were carried out as described by Sambrook et al. (1989). Radioactive probes were made using the Amersham Megaprime DNA labeling system (Amersham, Buckinghamshire, UK). Southern and colony hybridisations were performed as outlined in Sambrook et al. (1989). The plasmid pADAP is shown in Figure 6.

15 pADAP *Bam*HI library was constructed using a Sigma 'Gigapack'® III XL packaging extract, as specified by the manufacturer (Stratagene, California, USA).

Introduction of plasmid DNA into *E. coli* and *S. entomophila*

pLAFR3 based derivatives were introduced into *S. entomophila* by tripartite matings on solid media as described previously (Finnegan & Sheratt, 1982) using the pRK2013 helper
20 plasmid (Figorski & Helanski, 1979). pACYC184 and pBR322 based plasmids were electroporated into *E. coli* and *S. entomophila* strains, using a Biorad Gene Pulser (2µF, 2.5KV, and 200 abns) (Dower et al. 1988).

Mutagenesis

Transposon insertions were generated in recombinant plasmids using the mini-*Tn10* derivative 103 (kanamycin resistant) as described by Kleckner et al. (1991). Insertions were recombined into pADAP by transforming A1MO2 (refer to Table 1) with the
5 described construct. After growth in non-selective media, bacteria were screened for resistance to kanamycin and loss of the pLAFR3 tetracycline resistance marker.

Bioassay against *Costelytra zealandica* larvae

Infection of *C. zealandica* larvae was determined by a standard bioassay where the healthy larvae, collected from the field, were individually fed squares of carrot which had been
10 rolled in colonies of bacteria grown overnight on solid media (resulting in approximately 10^5 cells/carrot square). Twelve, second or third instar larvae were used for each treatment. Inoculated larvae were maintained at 15°C, in ice-cube trays. Larvae were left feeding on treated carrot for 3-4 days, then transferred to fresh trays and provided with untreated carrot for 10-14 days. The occurrence of gut clearance and loss of feeding was recorded every 3-4
15 days. Strains were considered disease-causing if greater than 70% of larvae showed disease symptoms by day 14. Known pathogenic and non pathogenic controls were included in all bioassays. Typically cessation of feeding occurs within 2-3 days while clearance of the larvae gut may take 4-6 days.

Recovery of bacteria from larvae

20 To isolate bacteria from inoculated grubs, larvae were surface sterilised by submerging in 70% methanol for 30 seconds. The larvae were then shaken in sterile DH_2O , removed and individually macerated in a 1.5ml microcentrifuge tube. The macerate was serial diluted and plated on LB media containing antibiotics selective for the host *S. entomophila* strain. To assess the stability of the bioassayed plasmid, colonies were patched onto a plate

containing antibiotics either selective for the recombinant plasmid or the *S. entomophilia* strain. Identity of plasmids in the recovered strain was checked by restriction enzyme profile.

Nucleotide Sequencing

- 5 A 9-kb *Bam*HI –*Eco*RI fragment derived from the pBM32-8 mutation (Fig 2b) and the 8kb *Hind*III fragment of pBM32 were separately cloned into the appropriate site of the deletion factory plasmid pDELTA1. Deletions were generated using the Deletion factory™ system (GIBCO BRL, MD, USA), as outlined in the manufacturers instructions. To identify the precise location of mini-*Tn*10 mutations, the peripheral mini-*Tn*10 *Bam*HI sites were used
- 10 in conjunction with the *Bam*HI sites of the pathogenic region to subclone the mini-*Tn*10 flanking regions into either pACYC184 or pUC19. Sequences were generated using the mini-*Tn*10 specific primer 5'ATGACAAGATGTGTATCCACC3' (Kleckner et al. 1991).

- Plasmids for sequencing were prepared by Wizard® (Promega, Madison, USA) or Quantum Prep® (Bio-Rad, California, USA) miniprep kits. Sequences were determined on both
- 15 strands, by using combinations of subcloned fragments, custom primers and deletion products derived from the deletion factory system (Gibco BRL, Madison, USA). The DNA was sequenced using either ³³P dCTP and the Thermosequenase cycle sequencing kit (Amersham, Buckinghamshire, UK), or by automated sequencing using an Applied Biosystem 373A or 377 autosequencer. Sequence data were assembled using SEQMAN
- 20 (DNASTAR Inc., Madison, USA). ORF's were analysed by Gene Jockey. Databases at the National Center for Biotechnology Information were searched by using BLASTN and BLASTX via the www.ncbi.nlm.gov/BLAST. Searches for DNA palindromes, repeats and inverted repeats were undertaken using DNAMAN (Lynnon Biosoft, Quebec, Canada). Protein motifs were searched using Blocks (<http://www.blocks.fhcrc.org/>), ExPASy
- 25 (<http://www.expasy.ch/>), and Gene Quiz (<http://columba.ebi.ac.uk:8765/gqsrv/submit>).

The sequences determined in this study have been deposited in Gene Bank under Accession Number AF1335182.

RESULTS

Cloning the disease encoding region from pADAP

5 Previously, Grkovic et al. (1995) have shown that the pADK-13 mutation can be complemented with the pADAP 11 kb *HindIII* fragment (pGLA-20). However, the pADK-10 mutation was unable to be complemented with pGLA-20. In an attempt to isolate the region that may complement the pADK-10 mutation the previously described pGLA-20 derived, pADK-35 null mutation (Grkovic et al. 1995) was used as a selective marker (Fig
10 1), to select the *BglIII* fragment encompassing both the pADK-10 and pADK-35 mutations. pADK-35 DNA was isolated and digested with the restriction enzyme *BglIII*. The resultant digest was ligated into the *BamHI* site of bBR322 to form the construct pBG35 (containing 12.8kb *BglIII*-mini-*Tn10* fragment). pBG35 was placed separately in *trans* with pADK-10 and pGLA-20, and the resultant strains bioassayed against grass grub larvae. Results
15 showed that pBG35 was able to complement the pADK-10 mutant, but was unable to induce any symptoms of amber disease when placed in *trans* with pGLA-20, indicating that there must be another region on pADAP needed to induce amber disease.

Restriction enzyme data of pGLA-20 and pBG35 suggested that the entire pathogenic region may reside within one of the large *BamHI* fragments of pADAP. A cosmid *BamHI*
20 library of pADAP was made and screened using the 2.2kb *EcoRI* fragment derived from pBG35 (Fig 1) as the probe. Several probe positive clones were isolated; all shared similar restriction enzyme profiles. However, one (designated pMH32) was found to be smaller, measuring only 23kb in size compared with the 33kb of the other clones (eg. pMH41; Fig 1b). The difference between pMH32 and pMH41 was found to be a 10kb deletion at the
25 left most end of pMH32 encompassing the one *HindIII* site (Fig 1). *E. coli* strains

containing pMH32 or pMH41 were bioassays against grass grub larvae and found to induce the full symptoms of amber disease (that is - gut clearance and antifeeding activity). However, about ten days after infection a proportion of grass grubs fed the *E. coli* strains were found to recover from a diseased to a healthy phenotype.

- 5 The plasmids pMH32 and pMH41 were subsequently introduced into a *S. entomophila* strain cured of pADAP (5.6RC) and the strains bioassayed against grass grub larvae. The strains gave the same disease progression as wild type and no larvae recovered, suggesting that the region cloned in pMH32 contained all the pathogenic determinants of pADAP.

Effect of copy number and mini-*Tn10* insertions in pBM32 on disease-causing ability

- 10 To facilitate mutagenesis and assess the effect of copy number on the disease process, the 23kb *Bam*HI fragment from pMH32 was cloned into the medium copy plasmid pBR322 to give pBM32. A bioassay comparing the ability of pMH32 and pBM32 to induce amber disease against grass grub was undertaken. Results showed that there were no visual differences in the progression of amber disease between pBM32 and pMH32. The
- 15 construct pBM32 was mutated with the mini-*Tn10* transposon derivative 103, and insertions mapped (Fig 2b). Bioassays of *E. coli* strains containing plasmids of the resultant mutants, showed that the disease determinants were confined within a central 16.9kb region (nucleotides 1955-18937 of SEQ ID NO: 1).

- All strains were non-pathogenic or fully pathogenic, and no partial disease phenotypes such
- 20 as antifeeding, or gut clearance were noted.

To confirm that no sequences at either end of the cloned fragment influenced the disease process, several deletion plasmids were made (Fig 2a). The large fragments resulting from cleavage of the pBM32 -4, -8, -10, -20, -23, -24 and -35 plasmids with *Bam*HI were cloned into the analogous site of pACYC184. The resultant plasmids were transformed into the

non-pathogenic *S. entomophilia* strain 5.6RKM and assessed for pathogenicity. This analysis confirmed that the central 16.9kb region (Fig 2a) was sufficient to induce the disease.

Effect of mini-*Tn10* insertions in pADAP on disease-causing ability

- 5 Grkovic et al. (1995) recombined by marker exchange the pGLA-20 based mutations - 10 and -13 into pADAP (Fig 2a). When bioassayed, *S. entomophilia* strains containing either of these mutant plasmids caused a partial condition including cessation of feeding but not gut clearance or amber colouration. This was in contrast to the complete abolition of disease observed in pADAP-cured *S. entomophilia* strains containing mutant pBM32
10 plasmids with similar insertions.

To determine the disease phenotype of the pBM32-based insertions in a pADAP background, the pBM32 based insertions were transferred into pADAP. pBM32 -1, -2, -4, -5, -6, -8, -9, -10, -21, -24, -30, -31 and -35 DNA fragments containing the inserted transposon and flanking DNA were cloned as independent fragments into pLAFR3 and the
15 inserts recombined back into pADAP by marker exchange (Fig 2c). The resultant recombinant *S. entomophilia* strains were checked by Southern analysis to confirm that recombination had occurred as expected and no pLAFR3 vector sequences were present (data not shown). Mutations that did not affect the disease process in pBM32 also had no effect when recombined back into pADAP. However, strains with the pADAP mutants that
20 totally abolished the disease process when in the pBM32 clone caused non-feeding but not gut clearance of the grubs (Fig 2b, c). Hence, none of the pADAP recombinant strains completely abolished the disease process. This suggests that, while the 16.9kb fragment contains all genes required for pathogenicity, other genes contributing to the antifeeding effect are present on some other part of pADAP.

- 25 Assessment of plasmid stability during the course of the bioassay showed that greater than

90% of the recombinant *Serratia* strains contained the clone of interest.

Nucleotide Sequence Analysis of the pathogenic region

- The large *Bam*HI fragment (18937 bp) derived from the pBM32-8 was sequenced on both strands using a combination of constructed detections, plasmid subclones and custom made
5 primers. A total continuous sequence of 18937 bp has been deposited in Gene Bank (Accession Number AF135182). Structural analysis of the DNA sequence using DNAMAN showed that there was a 12-bp sequence repeated five times between positions 683 and 743. The repeat is flanked by an upstream 13 base pair palindrome (669-682-bp), and a degenerate 34-bp downstream palindrome (765-799-bp)(Fig 2d,e).
- 10 Translation of the nucleotide sequence revealed nine significant open reading frames (ORF's). These together with their putative ribosomal binding sites and their base composition are listed in Table 2. Eight of the ORF's were oriented in the same direction and the other two in the opposite direction (Fig 2d). Sequence similarity searches showed that the deduced products of seven of these ORF's shared similarity with known proteins
15 (Table 3). Products of three of the ORF's showed similarity to different protein components of insecticidal toxins of *Photobacterium luminescens* (Bowen et al. 1998).

These ORF's have been designated *sep*. (*sepA*, *sepB* and *sepC*) for *Serratia entomophila* pathogenicity.

Similarities of deduced amino-acid sequences to proteins in current database

- 20 Results of database searches for homologues proteins are listed in Table 4.

With reference to Fig 2d and Table 4, the following protein similarities were identified:

The protein product of *sepA*, had high similarity to the *P. luminescens* insecticidal toxin complex protein TcbA, TcdA, TcaB and TccB. These proteins shared three significant

regions of predicted amino-acid similarity, at the amino-terminal region (SepA amino-acid residues (121-178), a central region (SepA amino-acid residues 960-1083) and, with greatest similarity, at the carboxyl terminus (SepA amino-acid residues 1630-2376) Fig. 4). However, there was little amino acid conservation around the putative proteolytic cleavage
5 site of TcaB, TcbA and TcdA identified by Bowen et al. (1998). SepA also contained a region (residues 1057-1345) with weak similarity to the *Clostridium bifermentans* mosquitocidal toxin cbm71 (Barloy et al., 1996).

SepB and the *P. luminescens* insecticidal toxin complex protein TcaC shared similarity throughout their length, and both SepA and TcaC showed high amino-terminal similarity to
10 the *Salmonella* virulence protein spvB (Gullig et al. 1992) (Fig. 5). The similarity of SepB and TcaC to SprB diminishes after SpvB amino acid residue 356.

SepC showed strong similarity to the amino-terminal of the insecticidal toxin complex protein TccC, up to amino-acid residue 663 of SepC. A number of putative bacterial cell wall proteins also have high similarity to SepC, including the wall associated protein
15 precursor *B. subtilis* (WAPA) and members of the *E. coli* Rhs (recombinant hot spot) elements. Strong similarity of SepC was also observed with hypothetical wall-associated proteins from *Coxiella burnetti* and *Bacillus subtilis* (Table 4).

The translated sequences of ORF1 and ORF2 showed no similarity to sequences in the current databases. ORF3 shared significant similarity to the morphogenesis protein of the
20 *Bacillus subtilis* bacteriophage B103, a member of bacteriophage muramidase-type lysis proteins (Pecenкова et al. 1996). However, relative to size, the gp19 protein of *S. typhimurium* phage ES18 (146 amino-acid residues) or the nucD/regB phage lysozymes of *S. marcescens* (179 amino-acid residues) are more similar. ORF4 showed similarity to *E. coli* bacteriophage N15gp 55 protein, a protein of unknown function (Zimmer et al. 1998).

25 Located in the same orientation as the sep genes and 134bp downstream of the *SepC*

termination codon is a 204 base pair region assigned ORF5, which has high similarity to a *S. typhimurium* resolvase/invertase protein. However ORF5 is disrupted by two stop codons at amino-acid residues 19 and 64, making it unlikely that an active resolvase/invertase protein, is encoded by this region. A 256-bp region of encompassed by
5 ORF5 (17498-17754) showed high similarity (77% identity) to the region (AF020806; 1629-1885 bp) encoding *S. typhimurium* DNA invertase gene (Valdivia et al. 1997) suggesting a similar ancestral origin.

Downstream of ORF5 and oriented in the opposite direction from 18935-18163 was a 870 base pair region of DNA designated ORF6 whose product showed high amino-acid
10 similarity over two different reading frames to the insertion element *IS91* of *E. coli* (Mendiola et al. 1992). The translated sequence is interrupted at amino-acid residue 149 of the *IS91* element and later resumed on a second reading frame, before its similarity switched back to the original reading frame. Switching of ORF's is a common feature of members of the IS3 family where the transposase is encoded by this overlapping ORF's
15 (Prere et al. 1990). However, the switch back to the initial strand is atypical. ORF6 may therefore be a dysfunctional relic of an ancestral *IS* element. It is unknown whether ORF6 contains a ribosomal binding site as its predicted location would lie outside the sequenced region. There was no DNA similarity to the *IS91* element.

Analysis for protein motifs showed that a tripeptide cell-binding motif Asp-Gly-Arg
20 (RGD), implicated in the binding of various adhesion proteins produced by parasites and viruses to eukaryotic cells (Leininger et al. 1991), is present in SepA and the *P. luminescens* TcdA, and TcaB proteins (Fig. 4). The RGD motif is present in cell surface adhesions produced by the human pathogen *Bordetella pertussis*, namely the filamentous haemagglutinin (220 kDa) (Relfman et al. 1989), and the outer membrane protein pertactin
25 (69 kDa) (Leininger et al. 1991). These motifs have been implicated in enhancing the binding of *B. pertussis* to eukaryotic cells. Because the RGD motif found in SepA falls in a

region of high similarity between SepA and its *P. luminescens* counterparts, it may play a role in mediating the attachment of the protein and/or the bacteria to the insect cell wall.

The hydropathicity profile of each of the Sep proteins was examined using the Kyte and Doolittle algorithm (Kyte and Doolittle, 1982) and compared to the relevant *P. luminescens* homologues. None of the Sep proteins contained a positively charged amino terminus followed by a hydrophobic region, characteristic of a signal sequence (Gierasch, 1989). The profiles of SepA, TcbA and TcdA were very similar (data not shown) and each exhibited a steep hydrophilic peak at the carboxyl terminus (residues 2055-2061 of SepA), specifically the protein sequence RRRRE (Fig. 4). Although both SepB and TcaC shared similarity to the *Salmonella* virulence protein SpvB, the amino-termini of SepB and TcaC were hydrophilic as opposed to the hydrophobic nature of SpvB. The profile of SepC and its *Photobacterium* counterpart TccC differed in that SepC had a slightly hydrophilic amino-terminus, whereas TccC lacked a hydrophilic amino-terminus and had a significantly hydrophobic carboxyl terminus from amino-acid residue 717 onwards (Fig. 3).

Analysis to detect repetitive motifs characteristic of the RTX family of toxins (Welch, 1991) using DOTPLOT showed only *P. luminescens* TccC contained a plot characteristic of a repeat motif present at the carboxy terminal (data not shown).

Analysis of DNA composition (%GC) and similarity

Comparisons of the GC content (Table 3) showed that the *SepA* and *SepB* genes were more GC-rich than their *P. luminescens* counterparts, while *SepC* and *tcaC* had similar GC content. The high GC content of *SepC* may be attributed to the close relationship of these protein products to the *rhs* family of wall-associated proteins which have a GC-rich core of 62% (Wang et al. 1998). Comparisons of the GC content of the *Sep* genes with that of the *S. entomophila* genome shows that they are rather similar, suggesting that the *sep* genes were not recently acquired by *S. entomophila*.

Identification of mini-*Tn10* location by sequence analysis

Analysis of the insertion points of the previously isolated mini-*Tn10* insertions (Fig. 2) within the putative ORF's (Table 4) revealed that ORF3 and ORF4 were interrupted by the -9, -23, -24 (ORF3) and -35 (ORF4) mutations. These insertions had no effect on the pathogenicity process, suggesting that ORF3 and ORF4 do not play a significant role in pathogenicity. However, the pADAP-35 mutation was at the 3' end of ORF4, resulting in the truncation of the final 11 amino-acid residues of ORF4 (Fig. 4), which may not have affected protein function. Further mutagenesis of ORF4 is therefore required to confirm that it has no role in pathogenicity. The mutations that caused loss of pathogenicity all resided within *SepA*, *SepB* or *SepC*. No mutation mapped to ORF1, ORF2 or ORF5.

Complementation analysis of the *sep* proteins

Following sequence data each of the *Sep* ORF's were excised as closely as possible with restriction enzymes, placed into pLAFR3 and placed in *trans* with the appropriate pADAP mutation. Complementation of *SepA* was undertaken through the use of the 8.5 kb *HindIII* clone (pMH45) which encompasses both ORF1 and *SepA*. *SepB* was excised as a 5.4 kb *StuI* fragment and *SepC* was excised as a 4.6 kb fragment using one of the peripheral; *BamHI* sites from the pBH32-13 mutation and the *StuI* site of pBM32 (Fig. 2b).

Complementation analysis showed that pLAFR3 based *SepB* and *SepC* are able to complement their mutated pADK- counterparts. Grkovic et al. (1995) had already previously shown that *SepC* could complement itself. However, this was achieved through using the entire 11 kb *HindIII*, pGLA-20 fragment.

Whether *SepA* is able to complement itself has yet to be fully established. It was found that ~98% of the pMH45 construct was lost during the course of the bioassay. This latter result was sporadic and occasionally a repeated experiment would show the presence of diseased

grubs. Analysis of the macerates of these grubs showed that pMH45 was present indicating that pMH45 can possibly complement *SepA*. However before further complementation analysis of *SepA* can be undertaken, measures to ensure the complementation plasmids stability are needed.

5 DISCUSSION

The large conjugative plasmid, pADAP, of *S. entomophila* encodes the genes responsible for cessation of feeding and gut clearance, characteristics of amber disease in the New Zealand grass grub *C. zealandica*. This plasmid is present in all *S. entomophila* and *S. proteamaculans* strains capable of causing amber disease (Glare et al. 1993) and had been
10 implicated in disease processes (Grkovic et al. 1995). The applicant has defined a 16.9 kb region of kADAP that is sufficient to confer pathogenicity towards *C. zealandica* on pADAP-cured strains of *S. entomophila* and on strains of *E. coli*. Hence, the region confers all the essential pathogenicity genes of *S. entomophila* responsible for amber disease. Nucleotide sequence and mutagenesis analysis of the region revealed three genes,
15 *SepA*, *SepB* and *SepC*, that together are sufficient for pathogenicity. Mutations in any of the three genes completely abolished the disease process and partial disease states were not detected, suggesting that the three genes may interact to exert an effect.

The 23-kb region cloned into pBR322 to make pBM32 conferred pathogenicity in pADAP-cured *S. entomophila* strains with all symptoms of amber disease being observed.
20 Insertion mutants in pBM32 that abolished pathogenicity were transferred to pADAP. The resultant strains showed a partial disease phenotype, including anti-feeding but not gut clearance, suggesting that an additional anti-feeding gene may be present elsewhere on pADAP. The occurrence of two different anti-feeding genes on pADAP also supports data of Grkovic et al. (1995) who found that suppression of feeding was stronger in the wild-
25 type pADK-6 strain, compared to the partial disease state (pADK-10, pADK-13) of

inducing anti-feeding but no gut clearance. A putative anti-feeding gene, *amb2*, has already been isolated from the genomic DNA of *S. entomophila* (Nunez-Valdez and Mahanty, 1996). Recent data indicate that the *amb2* locus resides at an as yet to be identified location on pADAP that is remote from the region identified herein (Hurst, unpublished data).

5 Sequence analysis and comparison of the products of the *sep* genes showed that they share significant similarity to the proteins TcbA (TcdA, TcaB, TccB), TcaC and TccC that comprise the toxin complexes of *P. luminescens*. Like the *P. luminescens* genes that *sep* genes of *Serratia* share a similar organisational pattern of three genes ordered in succession in the same orientation, and opposed by a terminal gene transcribed in the opposite
10 direction. However, the order of *sep* genes differ, are slightly smaller in size, and comprise constituents of each of the *P. luminescens* loci *tca* (*tcaB*=*sepA*, *tcaC*=*sepB*), *luminescens* toxin gene *tcd* (Ensign et al. 1997) is also similar to *SepA*. The similarity shared between the *sep* and *tc* gene products suggests that they are members of a new family of insecticidal toxins. The lack of DNA similarity as opposed to protein similarity between *sep* and *P.*
15 *luminescens* *tc* genes together with the difference in GC content of the *sepA* and *sepB* genes compared to the *tc* genes, suggests that these genes were present in the common enterobacterial ancestor of *P. luminescens* and *S. entomophila* and were not acquired by a more recent horizontal transfer event.

The *Photorhabdus* toxins were isolated as a composite of proteins which are hypothesised
20 to interact synergistically to form a toxin complex. The toxins are also able to exert an anti-feeding effect (Bowen et al. 1998; Bowen and Ensign, 1998). This is consistent with the results we obtained with the *sep* mutants. pADAP-cured *S. entomophila* strains containing the pathogenicity clone pBM32 exert an anti-feeding effect on the grass grub and individual mutations within any of the *sep* genes have an identical phenotype,
25 completely abolishing pathogenicity. The *Photorhabdus* toxins have a wide host range, affecting Lepidoptera, Coleoptera and Dictyoptera and undergo post translational

proteolytic processing (Bowen et al. 1998). No similarities of *sep* proteins were found to the *Photorhabdus* toxin component TccA, and only the amino-terminus of TcaA shared similarity to *SepA*. This and the difference in the hydrophobicity profiles of *SepC* and TccC, may account for specificity of the *sep* proteins towards *C. zealandica*. However the
5 *sep* proteins have yet to be purified and it is unknown whether the *sep* genes are expressed when *S. entomophila* is ingested by other insects. Therefore the possibility that these newly-described toxins may exhibit a broader host range cannot be ruled out.

The *Photorhabdus* toxin TcbA shares weak similarity to the *Clostridium difficile* A and B toxins (Bowen, 1998), but no such similarities were found to *SepA*. *C. difficile* A and B
10 toxins belong to the RTX (repeats in toxin) family of toxins which are noted for the presence of several carboxyl terminal repeats (von Eichel-Streiber et al. 1992). A search of the *sep* proteins and their *P. luminescens* homologues for protein repeats showed that only the *P. luminescens* TcaC protein contained a repeat-type signature. The TcaC carboxy-terminal repeat bears little resemblance in size or number of repeats found in RTX toxins
15 (von Eichel-Streiber et al. 1992). *SepA* does not show weak similarity to the mosquitocidal toxin Cbm71 of *C. bifermentans* (Barloy et al. 1996). However when this region is compared with the relevant *Photorhabdus* homologues, it is a region with little similarity.

SepB has strong similarities to both *P. luminescens* TccC and the *Salmonella* virulence gene product SpvB (Gulig et al. 1992). SpvB is believed to enhance the survival of virulent
20 *Salmonella* in macrophages (Libby et al. 1997). It has been suggested that TcaC may act by attacking insect haemocytes (Bowen et al. 1998). However, haemocytes reside within the insect haemocoel and *S. entomophila* does not invade the haemocoel until late in the infection process (Jackson et al. 1993), suggesting that *SepB* may act in some other way. The similarity of *SepB* and TcaC is high to SpvB but diminishes ten amino-acid residues
25 upstream of the proline-rich region found in SpvB that is postulated to divide the protein into separate domains (Roudier et al. 1992). This may indicate a vital role for the amino-

terminus of both *SepB* and *SpvB* in interacting with an evolutionarily-conserved eukaryotic protein.

The *SepC* protein shows high similarity to a family of cell wall-associated bacterial proteins such as the *B. subtilis* wall-associated protein (WAPA) and members of the *E. coli* rhs element family. The function of the Rhs proteins has yet to be established, but they are believed to be cell surface ligand-binding proteins (Hill et al. 1994). The Rhs proteins and the *B. subtilis* was-associated protein contain a characteristic repetitive peptide motif, but no such motif was observed in *SepC*. A feature of rhs elements is the presence of a downstream IS element (Wang et al. 1998). A degenerate IS91-type transposase element (ORF6) is present downstream of *SepC*. The IS91 element has been found associated with plasmids or chromosomal genes involved in α -haemolysin synthesis, and has been postulated to play a pivotal role in the spread of the α -haemolysin genes by means of the IS91-mediated recombinational activity (Zabala et al. 1984). It seems possible an IS element adjacent to *SepC* may have been involved in the acquisition of the *sep* genes by *S. entomophila*.

Blackburn et al. (1998) undertook histological examinations of the lepidopteran *Manduca sexta* after treatment with the *P. luminescens* Tca toxin complex introduced by feeding or haemcoelic injection. They found blebbing of the midgut epithelium into the lumen, resulting in lysis and formation of cavities. Similar histological studies have been undertaken at various stages throughout the infection cycle of *S. entomophila* in *C. zealandica*, and reveal a visible deterioration in the number of fat cells to almost minimal levels, and an emptying of the larval gut. However no blebbing of the midgut epithelium was observed (Jackson et al. 1993).

The *S. entomophila* pathogenicity region endows pathogenicity on members of the Enterobacteraceae such as *Klebsiella spp.*, *Enterobacter agglomerans*, *E. coli*, and *Serratia*

species (Glare et al. 1996). From this we can infer that the *Sep* proteins are the major virulence determinants, that the promoters of the *sep* genes are expressed constitutively or under the control of conserved regulatory genes, or a negative regulatory gene present in the pathogenicity region, and that export of the toxin proteins is carried out by a conserved

5 chromosomally encoded system, or is an intrinsic property of the *sep* proteins. The *Sep* proteins have no obvious amino terminal signal sequences, a facet shared with E-Group colicins. The release of cloacin DF13 is mediated through a small lipoprotein designated BRP, for bacteriocin-release protein. Low level expression of BRP in conjunction with phospholipase A leads to the release of cloacin DF13, along with bacterial periplasmic

10 proteins. However if expressed in high amounts, BRP causes cell death by cell lysis (vader Wal, 1998). The close proximity and similar orientation pattern of ORF3 to the *sep* genes indicate that ORF3 may have an as yet to be determined important functional role. Protein similarity searches show that it has high similarity to the bacteriophage lysozyme family. In relation to amino-acid size, ORF3 closely resembles the LZBP22 lysozyme of

15 the *Salmonella* P2 bacteriophage, a protein essential for the lysis of the bacterial cell wall (Rennell and Poteete, 1985). It is possible that ORF3 may facilitate the release of the *sep* proteins by lysing the bacterial cell wall. A low level expression of ORF3 might, as in the case of BRP, allow the passage of the *sep* proteins across the cell wall without causing cell death. The reason that the pBM32-9 and -24 mutations were unable to abolish the disease

20 process could be due to a masking of ORF3 function by natural cell lysis of the bacteria.

A region of repetitive DNA was identified between nucleotides 683 to 743, centered within a 1.2-kb AT rich stretch of DNA that contains no potential ORF's. The repeat motif is flanked by an upstream 13-bp palindrome and a degenerate downstream 33-bp palindrome. Repeats have been found to be common sites for recombination (Allgood et al. 1988), or to

25 facilitate the binding of proteins. A 66-bp DNA sequence termed the *rsk* element for reduced serum killing, of the *S. typhimurium* 95-kb virulence plasmid, comprises of a series

of direct 10-bp repeats with a 21 nucleotide periodicity. The *rsk* element is believed to titrate out a *trans*-acting factor, enhancing the expression of the *Salmonella* serum resistance gene (Vandenbosch et al. 1989). It is not known whether these repeats and/or flanking palindromes have a role in the pathogenicity process. The deletion derivative
5 pAC24, which encompasses this region, was still pathogenic towards the grass grub. However, this deletion could also unknowingly remove the complete regulatory circuit of the pathogenicity region, leading to constitutive expression.

THE ARABINOSE EXPRESSION SYSTEM

Methodology

10 Using the polymerase chain reaction (PCR) the initiation codon ATG of the three *sep* genes (*sepA*, *sepB* and *sepC*) were individually placed into the unique *NdeI* site (restriction enzyme site CATGG) of the HIS-tag arabinose expression vector pAV2-10 (obtained from Chuck Shoemaker –AgResearch). Because large proteins i.e. greater than 50 kda are limited in their ability to bind to HIS tag affinity columns the carboxyl terminus of each of
15 the Sep proteins did not need to be in frame with the HIS-tag site. Instead wild type DNA (non PCRd) containing a downstream chloramphenicol resistance gene was ligated into the appropriate restriction enzyme site (*sepA* *SunI*; *sepB* *HindIII*; *sepC* *BstXI*) of the pAV2-10-*sep* derived vectors:-

-the use of the chloramphenicol resistant marker provided by the vector pACYC184
20 enhances the stability to each of the expression constructs i.e. -the antibiotic ampicillin to which the pAV2-10 is resistant too is cleaved in the media to an inactive form leading to possible plasmid free segregants arising. Conversely the antibiotic chloramphenicol is not cleaved heightening the level of plasmid stability under conditions of arabinose induction.

To validate the legitimacy of the fused genes to the arabinose expression vector, PCR generated products and the ligation junctions were verified by DNA sequencing.

Concurrent to this the *sepB* and *sepC* genes were placed as derived from pADAP downstream of *sepA*. Also *sepA*, *sepB* and *sepC* were placed as in pADAP downstream of
5 orf3. This simulated wildtype conditions (i.e. the arrangement of the *sep* genes on pADAP) and hopefully get the production of the *sep* genes and the complex driven off the one upstream promoter. A method which Western analysis has shown to be successful –with moderate levels of *sepA*, *sepB* and *sepC* being detected.

The arabinose expression system is one of the tightest systems known with almost complete
10 abolition of gene product under arabinose free conditions Guzman *et al.* (1995), this abolition can be enhanced by providing glucose to the medium. In contrast providing arabinose at the concentration of 0.2% will switch the arabinose promoter on express any genes under its control e.g. *sepA* etc. Typically an overnight culture of the *E. coli* strain was set up the next day an 100 µl of the culture was suspended in fresh media
15 supplemented with chloramphenicol (30 µg/ml) the culture was grown until an OD of 400 at which time arabinose was added to the culture to a final concentration of 0.2% and the culture left shaking at 30 °C for 18 hours.

To date Western analysis has shown that each of the proteins is expressed and expressed to its correct predicted size:

20 SepA 262.7 kdal

 SepB 156.6 kdal

 SepC 107 kdal

SepC is expressed at high levels with minor levels of proteolytic cleavage. However both SepA and SepB though expressed are cleaved in high amounts by endogenous *E. coli* proteases. Alternative strains of *E. coli* are going to be assessed for loss of proteolytic activity against SepA and SepB

- 5 It has also been shown that placing all three of the *sep* genes under the control of a single arabinose promoter will result in the production of basal levels of the SepA, SepB, SepC toxin complex.

Each of the following Coleopteran species were mouth injected with 3-5 μ l of an overnight suspension of induced bacteria (*E. coli* strain DHB101) containing either SepA, SepB and

- 10 SepC or orf3, SepA, SepB and SepC.

Each larvae was then given a 3mm³ piece of carrot coated with a 50% solution (dH₂O) of arabinose. Observations were noted each day and the larvae refed with a 3mm³ piece of carrot coated with a 50% solution (dH₂O) of arabinose

Red headed cock chaffer

- 15 Tasmanian grass grub

Odontara

Grass grub (positive control)

- 20 Under these conditions it has been found that the arabinose expressed toxin complex SepA, SepB and SepC is active against grass grub but not any of the other species of scarabs tested (see above). It is therefore thought unlikely that the toxin complex will have activity to other insect orders.

SUMMARY

The bacteria *Serratia entomophila* and *S. proteamaculans* cause amber disease in the grass grub, *Costelytra zealandica* (Coleoptera: Scarabaeidae), an important pasture pest in New Zealand. Larval disease symptoms include amber colouration, clearance of the gut and
5 rapid cessation of feeding, before eventual death. The region containing pathogenic determinants of the disease has been cloned, and further defined by mutagenesis and deletion analysis to a 16.9 kb region. Sequence analysis of the minimal pathogenic encoding region showed significant protein homology, but little sequence homology to a group of newly described toxins from a member of the Enterobacteriaceae, *Photorhabdus*
10 *luminescens*. This pathogenicity-encoding region from *S. entomophila* plasmid pADAP is the subject of the invention. The proteins encoded by the genes (*sepA*, *sepB*, *sepC*) within the 16.9 kb region can be used for insect control whether as an inundative pesticide, within baits or expressed in other organisms such as plants or microbes.

Aspects of the present invention have been described by way of example only and it should
15 be appreciated that modifications and additions may be made thereto without departing from the scope thereof as defined in the appended claims.

REFERENCES

- Barloy F; Delecluse A; Nicolas L and Lecadet M M (1996)
Cloning and expression of the first anaerobic toxic gene from *Clostridium bifermentans* subsp. *malaysia*, encoding a new mosquitocidal protein with homologues to *Bacillus thuringiensis* delta-endotoxins. *J. Bact.* 178 : 3099-3105.
- Blackburn M; Golubeva E; Bowen D and Ffrench-Constant R H (1998)
A novel insecticidal toxin from *Photorhabdus luminescens*, Toxin complex a (*Tca*), and Its Histopathological Effects on the Midgut of *Manduca sexta*. *Applied and Environmental Microbiology* 64: pp 3036-3041.
- Bolivar F; Rodriguez R L; Greene P J; Betlach M C; Heyneker H L and Boyer H W (1977)
Construction and characterisation of new cloning vehicles II. A multipurpose cloning system. *Gene* 2: 95-113.
- Bowen D J and Ensign J C (1998)
Purification and characterisation of a High-Molecular-Weight Insecticidal Protein Complex produced by the Entomopathogenic Bacterium *Photorhabdus luminescens* *Applied and Environmental Microbiology* 64: pp 3029-3035.
- Bowen D; Rocheleau; Blackburn M; Andreev O; Golubeva E; Bharia R and Ffrench-Constant R H (1998)
Insecticidal Toxins from the Bacterium *Photorhabdus luminescens* *Science* 280: pp 2129-2132.
- Casabadian M J and Cohen S N (1980)
Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.* 138 : 179-207.
- Chang A C Y and Cohen S N (1978)
Construction and characterisation of amplifiable multicopy DNA cloning vehicles derived from the p15A cryptic miniplasmid. *J. Bact* 134(3) : 1141-1156.
- Corbett (unpublished)

- Ditta G; Stanfield S; Corbin D and Helinski D R (1980)
Broad host range cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA. 27 : 7347-7351
- Dower W J; Miller J F and Ragsdale C W (1988)
High efficacy transformation of *E.coli* by high voltage electroporation. Nucleic Acids Res. 16 : 6127-6145.
- Figurski D H and Helinski D R (1979)
Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proc. Natl. Acad. Sci. USA. 76 : 1648-1652.
- Finnegan J and Sherrat D (1982) Plasmid ColE1 conjugal mobility: the nature of *bom*, a region required in *cis* for transfer. Mol. Gen. Genet. 185, 344-351.
- Gierasch, L M (1989) Signal sequences. *Biochem* 28: 923-930
- Glare T R; Corbett G E and Sadler A J (1993)
Association of a large plasmid with amber disease of the New Zealand grass grub, *Costelytra zealandica*, caused by *Serratia entomophila* and *Serratia proteamaculans*. Journal of Invertebrate Pathology 62, 165-170.
- Glare T R; Hurst M R H and Grkovic S (1996)
Plasmid transfer among several members of the family Enterobacteriaceae increases the number of species capable of causing experimental amber disease in grass grub. FEMS Microbiology Letters 139: 117-120.
- Grimont P A D; Jackson T A; Ageron E and Noonan M J (1988)
Serratia entomophila sp. nov. associated with amber disease in the New Zealand grass grub, *Costelytra zealandica* Int. J. System. Bacteriol, 38 : 1-6.
- Grkovic S; Glare T R; Jackson T A and Corbett G E (1995)
Genes essential for amber disease in grass grub are located on the large plasmid found in *Serratia entomophila* and *Serratia proteamaculans*. Applied and Environmental Microbiology 61, 2218-2223.

Gulig P A; Caldwell A L and Chiodo V A (1992)

Identification, genetic analysis and DNA sequence of a 7.8-kb virulence region of the *Salmonella typhimurium* virulence plasmid. *Mol. Microbiol.* 6 : 1395-1411.

Hanahan D (1983)

Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166 : 557.

Jackson T A; Huger A M and Glare T R (1993)

Pathology of amber disease in the New Zealand grass grub, *Costelytra zealandica* (Coleoptera: Scarabaeidae). *J Invertebr. Pathol.*, 61: 123-130.

Jackson T A (1995)

Amber disease reduces trypsin activity in midgut of *Costelytra zealandica* larvae. *J. Invert. Pathol.* 65: 68-69.

Kleckner N; Bender J and Gottesman S (1991)

Uses of transposons with emphasis on Tn10. *Methods Enzymol* 204: 139-179.

Kyte J and Doolittle R F (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157: 105-132

Leininger E, Roberts M, Kenimer J G, Charles IG, Fairweather N, Novotny P, and Brennan M'J (1991) Pertactin, an Arg-Gly-Asp-containing *Bordetella pertussis* surface protein that promotes adherence to mammalian cells. *Proc Natl Acad Sci USA* 88: 345-349

Lorrow D and Jesse J (1990)

Max efficiency DH10B™: A host for cloning methylated DNA. *Focus* 12: 19.

Mendiola M V; Jubete Y and de la Cruz F (1992)

DNA sequence of IS91 and identification of the Transposase Gene. *Journal of Bacteriology* 174: 1345-1351.

Nunez-Valdez M E and Mahanty H K (1996)

The *amb2* locus from *Serratia entomophila* confers anti-feeding effect on larvae of *Costelytra zealandica* (Coleoptera: Scarabaeidae). *Gene* 172: 75-79.

- Pecenкова T; Benes V; Paces J; Vlcek C and Paces V (1996)
Bacteriophage B103: complete DNA sequence of its genome and relationship to other *Bacillus* phages. *Gene* 199 157-163.
- Prere M F, Chandler M and Fayet O (1990) Transposition in *Shigella dysenteriae* isolation and analysis of IS911, a new member of the IS3 group of insertion sequences. *J Bacteriol* 172: 4090-4099.
- Relman D A, Domenighini M, Tuomanen E, Rappuoli R and Falkow S (1989)
Filamentous hemagglutinin of *Bordetella pertussis*: nucleotide sequence and crucial role in adherence. *Proc Natl Acad Sci USA* 86: 2637-2641.
- Sambrook J; Fritsch E F and Maniatis T (1989)
Molecular cloning, 2nd edition, Cold Springs Harbour Laboratory Press, Cold Spring Harbour
- Staskawicz B; Dahlbeck D; Keen N and Napoli C (1987)
Molecular characterization of cloned avirulence genes from Race 0 to Race 1 of *Pseudomonas syringae* pv *slycinea*. *J. Bacteriol* 169 : 5789-5794.
- Stucki G; Jackson T A and Noonan M J (1984)
Isolation and characterisation of *Serratia* strains pathogenic for larvae of the New Zealand grass grub *Costelytra zealandica*. *NZ J Science* 27: 255-260.
- Trought T E T; Jackson T A and French R A (1982)
Incidence and transmission of a disease of grass grub (*Costelytra zealandica*) in Canterbury. *NZ J. Exp. Agric.* 10: 79-82.
- Upadhyaya N M; Glare T R and Mahanty H K (1992)
Identification of a *Serratia entomophila* genetic locus encoding amber disease in New Zealand grass grub (*Costelytra zealandica*). *J. Bacteriol* 174: 1020-1028.
- Valdivida R H and Falkow S (1997) Fluorescence-based isolation of bacterial genes expressed within host cells. *Science* 277 (5334), 2007-2011.
- Wang Y D; Zhao S and Hill C H (1998)
Rhs elements comprise three subfamilies which diverged prior to acquisition by *Escherichia coli* *J. Bacteriol.* 180 : 4102-10.

Welch R A (1991)

Pore-forming cytolysins of Gram-negative bacteria. *Mol. Microbiol* 5 : 521-528.

Yanisch-Perron C; Vieira J and Messing J (1985)

Improved M13 phage cloning vectors and host strains: nucleotide sequence of M13mp18 and pUC19 vectors. *Gene* 33, 103-119.

Zimmer A and Schmieger H.

Lysis gene modules in the phage P22 gene pool Zimmer A; Institute for Genetics and Microbiology, University of Munich, Maria-Ward-Str. 1a, Muenchen D-80638, Germany X167137. Accession number (AF064539).

Guzman L-M., Belin, D., Carson, M.J., and Beckwith, J. (1995): Tight regulation, modulation , and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J Bacteriol.* 177: 4121-4130.

Table 1 Bacterial strains, plasmids and bacteriophage used in the study

Bacteria	Description	Reference
<i>Escherichia coli</i>		
DH5 α	F ϕ 80d <i>lacZ</i> pM15 p(<i>lacZYA-argF</i>)U169 <i>recA1 endA1 supE44</i>	Hanahan (1983)
DH10B	F <i>mcrA</i> p(<i>mrr-hsdRMS-mcrBC</i>) ϕ 80d <i>lacZ</i> pM15 <i>placX74 endA1 recA1 deoRp(ara, leu)</i> 7697 <i>araD139 galU galK nupG rpsL λ</i> .	Lorow and Jessee, (1990)
DF1	$\gamma\delta$ transposase(<i>tnpA</i>)	Gibco BRL
MC1061	<i>sup^o hsdR mcrB araD139 p(araA BC-leu)</i> 7679 <i>placX74 galU galK rpsL thi</i>	Casadaban and Cohen, (1980)
MC4100	<i>araD139 p(lacZYA-argF)</i> U169 <i>rpsL150 St^R relA1 flbB5301 deoC1 ptsF25 rbsR</i>	Silhavy <i>et al.</i> (1984)
XL1-BlueMRA	p(<i>mcrA</i>)183 p(<i>mcrCB-hsdSMR-mrr</i>)173 <i>endA1 supE44 thi-1 reA1 gyrA96 relA1</i>	Stratagene
<i>Serratia entomophila</i>		
A1MO2	Ap ^R , pADAP, pathogenic.	Grimont <i>et al.</i> (1988)
5.6	heat cured pADAP minus derivative of A1MO2	Glare <i>et al.</i> (1993)
5.6RC	Cm ^R <i>recA</i> ⁻ pADAP minus strain	Grkovic <i>et al.</i> (1996)
5.6RK	Kn ^R <i>recA</i> ⁻ pADAP minus strain	this study
Plasmids		
pACYC184	Cm ^R Tc ^R	Chang and Cohen, (1978)
pADAP	Amber disease associated plasmid	Glare <i>et al.</i> 1993)
pBR322	Ap ^R , Tc ^R	Bolivar <i>et al.</i> (1977)
pBM32	23-kb <i>Bam</i> HI fragment from pMH32 cloned in pBR322	this study
pBM32-1-40	pBM32 containing mini- <i>Tn10</i> insertions	Gibco BRL
pDELTA1	Ap ^R , Sm ^R , Kn ^R , sucrose ^R	Staskawicz <i>et al.</i> (1987)
pLAFR3	Tc ^R pRK290 with λ cos, <i>lacZα</i> and multi-cloning site from pUC8.	Ditta <i>et al.</i> (1980)
pRK2013	IncP, Kn ^R Tra RK2 <i>repRK2 repE1</i>	Corbett (unpublished)
pGLA20	10.6-kb <i>Hind</i> III pADAP fragment cloned in pLAFR3	
pACp4	19-kb <i>Bam</i> HI fragment from pBM32-4 cloned in pACYC184	this study
pACp8	17-kb <i>Bam</i> HI fragment from pBM32-8 cloned in pACYC184	this study
pACp10	19.5-kb <i>Bam</i> HI fragment from pBM32-10 cloned in pACYC184	this study
pACp20	20-kb <i>Bam</i> HI fragment from pBM32-20 cloned in pACYC184	this study
pACp23	21-kb <i>Bam</i> HI fragment from pBM32-23 cloned in pACYC184	this study
pACp24	21.2-kb <i>Bam</i> HI fragment from pBM32-24 cloned in pACYC184	this study
pADK-10	pADAP::mini- <i>Tn10</i> insertion in 10.6-kb <i>Hind</i> III fragment, Kn ^R non-pathogenic	Grkovic <i>et al.</i> (1995)
pADK-13	pADAP::mini- <i>Tn10</i> insertion in 10.6-kb <i>Hind</i> III fragment, Kn ^R non-pathogenic	Grkovic <i>et al.</i> (1995)
pADK-35	pADAP::mini- <i>Tn10</i> insertion in 10.6-kb <i>Hind</i> III	Grkovic <i>et al.</i> (1995)

pMH32	fragment, Kn^R , pathogenic 23-kb <i>Bam</i> HI fragment of pADAP cloned into pLAFR3	this study
pMH41	33-kb <i>Bam</i> HI fragment of pADAP cloned into pLAFR3	this study
pBM32	23-kb <i>Bam</i> HI fragment of pMH32 cloned into pBR322	this study
pUC19	Ap^R , <i>lacZ</i> α , multi-cloning site	Yannish-Perron, <i>et al.</i> (1985)
Bacteriophage		
λ NK1316	mini-Tn10 derivative 103 donor λ b522 c1857 Pam80 nin5	Kleckner <i>et al.</i> (1991)

Table 2 Position of genes and features of the predicted gene products encoded by *sep* genes

ORF	Putative ribosome-binding site ^a	Longest potential coding region		<i>sep</i> %GC (<i>P. luminescens</i> homologue, %GC)
		Start at nucleotide	Stop at nt (ORF size bp)	
<i>sepA</i>	ATGGG <u>AC</u> CAACGTAATGAA TGAGG	2413	9547 (7131)	54 (<i>tcbA</i> , 43; <i>tcdA</i> , 44)
<i>sepB</i>	CGAGGAG <u>ACT</u> GAGCATGCAA	9598	13885 (4287)	58 (<i>tcaC</i> , 51)
<i>sepC</i>	ACAGGAGATCACATGAGC	14545	17467 (2922)	55 (<i>tccC</i> , 54)
ORF1	CATAGAG <u>ACT</u> GTGCTATGTTA	1287	1587 (300)	39
ORF2	TTGGAG <u>A</u> ATAACCGCCATGTT	1590	1863 (273)	39
ORF3	GGGGGAGAAAAATGAAG	1860	2294 (435)	51
ORF4	TGACTGGGAAGGAGGGGGGAC GGTGATGAGT	13908	14483 (576)	60
ORF5	TAACGAG <u>ACT</u> TTTTAGCAAAAT GGCACTTT	1761-1755, 1755-1773		?
ORF6	GAGCATGGC-Mini-Tn10-8*	18934-18064		?

* Putative ribosome-binding sites are underlined, and potential start codons are in boldface; nt, nucleotides; ? degenerate or incomplete ORF. * ORF transcribed in opposing direction.

Table 3. Comparisons of GC content between the *Sep* and *P. luminescens* genes

<i>Sep</i> (%GC)	<i>P. luminescens</i> toxin (%GC)
<i>sepA</i> (54%)	<i>tcbA</i> (43%) <i>tcdA</i> (44%)
<i>sepB</i> (58%)	<i>tcaC</i> (51%)
<i>sepC</i> (55%)	<i>tccC</i> (54%)

Table 4. Similarities of products of putative ORF's to protein sequences in the database detected using BlastP

ORF (a.a size)	Protein homo- logue (a.a size)	Degree of similarity %identity/%similarity (over) a.a residue - a.a residue	Function of the homologous protein	Organism	Blast score Reference ^a
SepA (2373)	TcbA (2504)	34/50 (1675) 41-1628* 57/72 (751) 1630-2374*	insecticidal toxin complex protein	<i>Photobacterium luminescens</i>	0.0 AF047457
	TcdA (2405)	40/55 (2458)*	insecticidal toxin complex protein	<i>P. luminescens</i>	0.0 Ensign <i>et al.</i> , (1997)
	TcaB (1189)	38/54 (764) 1625-2374* 29/50 (281) 936-1198*	insecticidal toxin complex protein	<i>P. luminescens</i>	e^{-137} AF046867
	TccB (1565)	36/51 (859) 1575-2373* 31/51 (289) 930-1204*	insecticidal toxin complex protein	<i>P. luminescens</i>	e^{-136} AF047028
	TcaA (1095)	36/56 (90) 94-183* 18/39 (530) 435-928*	insecticidal toxin complex protein	<i>P. luminescens</i>	$1e^{-4}$ AF046867
	TccA (965)	27/45 (186) 115-280*	insecticidal toxin complex protein	<i>P. luminescens</i>	$5e^{-4}$ AF047028
	Cbm71 (613)	24/41 (199) 1057-1250*	Mosquitocidal toxin Cbm71	<i>Clostridium bifermentans</i>	g2127309
SepB (1428)	TcaC (1485)	49/63 (1276) 1-1263* 64/78 (152) 1270-1421*	insecticidal toxin complex protein	<i>P. luminescens</i>	0.0 AF046867
	SpvB (591)	40/52 (357) 9-365*	<i>Salmonella</i> virulence protein	<i>Salmonella typhimurium</i>	$4e^{-52}$ S22664
SepC (938)	TccC (1043)	53/66 (836) 3-782*	insecticidal toxin complex protein	<i>P. luminescens</i>	0.0 AF047028
	SC2H4.02 (2183)	23/34 (639) 68-677*	Hypothetical wall associated protein	<i>Streptomyces coelicolor</i>	$2e^{-12}$ AL031514.1
	WapA (2334)	22/34 (430) 255-677* 20/36 (613) 48-625*	Wall associated protein Precursor	<i>B. subtilis</i>	$2e^{-5}$ S32920
	Y15898 (334)	21/34 (542) 181-684*	hypothetical wall associated protein	<i>Coxiella burnetii</i>	$9e^{-5}$ Y15898
	Rhs core (1420)	21/35 (463) 237-677* 21/36 (285) 35-300*	Rhs core protein	<i>E. coli</i>	$3e^{-4}$ AF044501
ORF3 (144)	BB103G (263)	45/62 (142) 1-139*	morphogenesis protein of bacteriophage B103	<i>Bacillus subtilis</i>	$3e^{-27}$ CAA67646
	LZBP22 (146)	46/61 (139) 1-143	Phage P22, lysozyme (E 3.2.1.17)	<i>Salmonella</i>	$1e^{-24}$ gi 138699
ORF4 (191)	Gp55 (181)	28/42 (188) 1-184*	bacteriophage N15 protein	<i>E. coli</i>	$1e^{-4}$ AF064539
ORF5 (236)	SprA	75/79(68) 1-68 ♦	Resolvase/invertase homologue	<i>S. typhimurium</i>	$7e^{-19}$ AF029069 AF020806
ORF6 (310)	IS91	39/56 (94) 130-197♦ -1* 39/58 (94) 224-318♦ -2* 30/48 (76) 319-395♦ -1*	IS91 transposase	<i>E. coli</i>	$4e^{-28}$ S23782

Percent identities and similarities were calculated in relation to the deduced gene products of the sequenced ORF. * indicates position of amino-acid similarity in relation to sequence generated in this study. ♦ indicates position of amino-acid similarity in relation to data base protein sequence. * reading frame. ^a similarities were considered potentially significant if the BlastP score exceeded e^{-3} .

Table 5 Positions of mini-Tn10 insertions

Mini-Tn10 insertion #	ORF	Position downstream of initiation codon (bp)
9/23	ORF3	120
24	ORF3	345
4	<i>sepA</i>	747
27	<i>sepA</i>	1037
40	<i>sepA</i>	1097
6	<i>sepA</i>	1727
38	<i>sepA</i>	2887
2	<i>sepA</i>	3197
5	<i>sepA</i>	3737
3	<i>sepA</i>	3697
19	<i>sepA</i>	3697
30	<i>sepA</i>	4467
37	<i>sepA</i>	4467
31	<i>sepA</i>	4627
12	<i>sepB</i>	182
22	<i>sepB</i>	172
11	<i>sepB</i>	362
10	<i>sepB</i>	2162
35	ORF4	557
13	<i>sepC</i>	2525
8		18937
ORF4/-35 junction GGG CGC <u>TGA</u> <u>TGA</u> ATC		

THE CLAIMS DEFINING THE INVENTION ARE:

1. A purified and isolated nucleic acid molecule comprising a nucleotide sequence of SEQ ID NO: 1 that encodes at least one of:
 - (i) an insecticidal protein complex, or
 - (ii) a functional fragment of said complex, or
 - (iii) a neutral mutation of said complex, or
 - (iv) a homolog of said complex,each of which have at least 75% nucleic acid homology to SEQ ID NO: 1 and are capable of hybridising with said nucleic acid molecule under stringent hybridisation conditions.
2. A purified and isolated nucleic acid molecule as claimed in Claim 1 comprising the nucleotide sequence 1995-18937 of SEQ ID NO: 1.
3. A purified and isolated nucleic acid molecule as claimed in Claim 1 comprising one or more of the nucleotide sequences 2411-9547, 9589-13883 or 14546-17467 of SEQ ID NO: 1.
4. A purified and isolated nucleic acid molecule as claimed in Claim 3 comprising all of nucleotide sequences 2411-9547, 9598-13884 and 14546-17467 of SEQ ID NO: 1.
5. A purified and isolated nucleic acid molecule as claimed in Claim 1 comprising a sequence of SEQ ID NO: 1, operably linked to at least one further nucleotide sequence which encode an insecticidal protein.
6. A purified and isolated nucleic acid molecule as claimed in Claim 2 comprising nucleotides 1955-18937 of SEQ ID NO: 1, operably linked to at least one further nucleotide sequence which encode an insecticidal protein.

7. A purified and isolated nucleic acid molecule as claimed in Claim 3 comprising a sequence of SEQ ID NO: 1, or one or more of nucleotides 2411-9547, 9598-13884 or 14546-17467 of SEQ ID NO: 1, operably linked to at least one further nucleotide sequence which encode an insecticidal protein.
8. A purified and isolated nucleic acid molecule as claimed in any one of claims 4 through 6 wherein the said nucleotide sequence includes the nucleotide sequence which codes for at least one of the *Bacillus* delta endo toxins, vegetative insecticidal proteins (vips), cholesterol oxidases, *Clostridium bifermentens* mosquitocidal toxins and/or *Photobacterium luminescens* toxins.
9. A purified and isolated nucleic acid molecule as claimed in claim 1 wherein nucleic acid molecule may comprise DNA, cDNA or RNA.
10. A purified and isolated nucleic acid molecule as claimed in claim 1 wherein the nucleic acid molecules said fragment, neutral mutation or homolog thereof capable of hybridising to said nucleic acid molecule, hybridise to the nucleotide sequence of SEQ ID NO: 1, or nucleotides 1955-18937, 2411-9547, 9598-13884 or 14546-17467 of SEQ ID NO: 1 if there is at least 75% or greater identity between the sequences.
11. A purified and isolated nucleic acid molecule as claimed in claim 1 wherein the nucleic acid molecule may be isolated from *Serratia entomophila* or *Serratia proteamaculans* strains of bacteria.
12. A recombinant expression vector(s) containing the nucleic acid molecule as claimed in Claim 1 and host transformed with the vector expressing a polypeptide.
13. A recombinant expression vector(s) as claimed in claim 11 wherein the vector is selectable from any suitable natural or artificial plasmid/vector.
14. A recombinant expression vector(s) as claimed in claim 13 wherein said suitable natural or

artificial plasmid/vector, including, pUC 19 (Yannish-Perron et al. 1995), pProEX HT (GibcoBRL, Gaithersburg, MD, USA), pBR322 (Bolivar et al. 1977), pACYC184 (Chang et al. 1978), pLAFR3 (Staskowicz et al. 1987).

15. A polypeptide resulting from the transformation or transfection of a host cell with a recombinant expression vector as claimed in any one of Claims 12 through 14.
16. A method of producing a polypeptide of claim 15 comprising the steps of:
 - (a) culturing a host cell which has been transformed or transfected with said vector as defined above to express the encoded polypeptide or peptide; and
 - (b) recovering the expressed polypeptide or peptide.
17. The use of a ligand that binds to a polypeptide of claim 15 to isolate and/or identify the polypeptide of claim 15.
18. An antibody or antibody binding fragment that binds to a polypeptide of claim 15.
19. Probes and primers comprising a fragment of the nucleic acid molecule as claimed in Claim 1 wherein said fragment is hybridisable under stringent conditions to a native insecticidal gene sequence.
20. Probes and primers comprising a fragment of the nucleic acid molecule as claimed in claim 19 wherein said probes and primers enable the structure and function of the gene to be determined and homologs of the gene to be obtained from bacteria other than *Serratia* sp.
21. A polypeptide as claimed in Claim 15 wherein the polypeptide has insecticidal activity encoded by the nucleic acid molecule of claim 1, or a functional fragment, neutral mutation or homolog thereof.
22. A polypeptide having insecticidal activity as claimed in claim 21 wherein the polypeptide

comprises the amino acid sequence of SEQ ID NO: 1 or a functional fragment, neutral mutation or homolog thereof.

23. A polypeptide having insecticidal activity as claimed in claim 21 wherein the polypeptide comprises amino acids 32-5118 of SEQ ID NO: 1.
24. A polypeptide having insecticidal activity as claimed in claim 21 wherein the polypeptide comprises at least one amino acid sequence of SEQ ID NO: 2; SEQ ID NO: 3; SEQ ID NO: 4; SEQ ID NO: 5 or SEQ ID NO: 6.
25. A polypeptide having insecticidal activity as claimed in claim 24 wherein the polypeptide preferably comprises amino acid sequence SEQ ID NO: 4; SEQ ID NO: 5 and SEQ ID NO: 6.
26. A polypeptide having insecticidal activity as claimed in claim 24 wherein the polypeptide preferably comprises all of SEQ ID NOs: 2-6.
27. A polypeptide having insecticidal activity as claimed in claim 21 wherein the polypeptide is obtained by expression of a DNA sequence coding therefore in a host cell or organism.
28. A polypeptide having insecticidal activity as claimed in claim 27 wherein the polypeptide comprises the amino acid sequence of SEQ ID NO: 1 linked to at least one further amino acid sequence encoding an insecticidal protein.
29. A polypeptide having insecticidal activity as claimed in claim 28 wherein the at least one further amino acid sequence includes the amino acid sequence which codes for *Bacillus* delta endo toxins, vegetative insecticidal proteins (vips), cholesterol oxidases, *Clostridium bifermentans* mosquitocidal toxins and/or *Photobacterium luminescens* toxins.
30. A polypeptide having insecticidal activity as claimed in claim 28 wherein the polypeptides comprise at least 50%, preferably 60%, more preferably 70% and most preferably 90-95% or greater identity to SEQ ID NO: 1.

31. A polypeptide having insecticidal activity as claimed in claim 21 wherein the polypeptide is produced by expression of a vector comprising the nucleic acid of SEQ ID No:1 or a functional fragment, neutral mutation or homolog thereof, in a suitable host cell.
32. An insecticidal composition comprising at least the polypeptide as claimed in claim 21 and an agriculturally acceptable carrier.
33. An insecticidal composition as claimed in claim 32 wherein more than one polypeptide is included in the composition.
34. An insecticidal composition as claimed in claim 32 or 33 wherein the composition comprises additional pesticides, including compounds known to possess herbicidal, fungicidal, insecticidal or nematocidal activity.
35. An insecticidal composition as claimed in claim 34 wherein the composition comprises other known insecticidally active agents, including *Bacillus delta* endo toxins, vegetative insecticidal proteins (vips), cholesterol oxidases, *Clostridium bifermentens* mosquitocidal toxins and/or *Photobacterium luminescens* toxins.
36. A method of combating pests, said method comprising applying to a locus, host and/or the pest, an effective amount of the polypeptide as claimed in Claim 21 that has functional insecticidal activity against said pest.
37. A method of inducing amber disease or like condition in insects comprising delivery to an insect an effective amount of the polypeptide as claimed in Claim 21 that has functional insecticidal activity against said insect.
38. A method of inducing amber disease or like condition in insects as claimed in claim 37 comprising delivery to an insect an effective amount of the polypeptide wherein the insect is selected from the order comprising Coleoptera.
39. A method of inducing amber disease or like condition in insects as claimed in Claim 38

comprising delivery to an insect an effective amount of the polypeptide wherein the insect includes *Costelytra zealandica* (Coleoptera: Scarabaeidae).

40. A method of delivering the insecticidal polypeptide to induce amber disease or like condition in insects including delivery of the insecticidal polypeptide as claimed in Claim 39 to the insect by any one of presenting the insecticidal polypeptide orally as a solid bait matrix, as a sprayable insecticide sprayed onto a substrate upon which the insect feeds, applied directly to the soil subsurface or as a drench or is expressed in an transgenic plant, bacterium, virus or fungus upon which the insect feeds.
41. A transgenic plant, bacterium virus or fungus, incorporating in its genome, a nucleic acid molecule as claimed in Claim 1 for providing the plant, bacterium virus or fungus with an ability to express an effective amount of an insecticidal polypeptide.

WO 01/16305

PCT/NZ00/00174

3

ccgtagcta taaatatgat atttaaactc gtatttttat ataaaaccag tttatgatgc 1080
 tggattggtc attaaagtcg ttatatgtga tcgttatctg tcattgattg gtgtttaatc 1140
 ttttattctt ccagtgaggt ttcaggggga atgtattggg taatcatact catgtcattt 1200
 gttgctttga tgttaaatta acgtgttcat tcattatggt ctactgttgt ttctattgtc 1260
 cggaacgacc atagagactg tcgctatggt aataggaata ttgactggt tatatgcgcc 1320
 aagggttatc gctgcactct ctggggcgat ggtattcatc attacgcaag ataacttcat 1380
 tgggtgcaga cgggtgttat tgtttttgt gtctttttta ctcggttga cattttcaga 1440
 gacaacagct tccgttatca acttctatat cccgaatgat atacatatag gaaatgacct 1500
 tgggtgccttt gttaccagcg ccgtgacggt gaagcttttt gttatcatta tgagcaagat 1560
 agagagaaaa tatcttgag aataaccgcc atgttccaaa tcatacttct taatgttaat 1620
 gccgtgattt gcttggtat tgccgtcaga ttattcctgt gccgtatcaa tcataaaatg 1680
 aaaaacattg tcgtctcttt tattgctttt ctctattatta cggcgtgcgg cgtgtctcc 1740
 atcaggacga tgacggggga gtattactat gccgattggt ccgagacgat cattaacctt 1800
 tcgcttttcc tgtctgttta tatacgcaat gccgaaatcc ttcggtgggg ggagaaaaa 1859
 atg aag ata agt tcc cga ggt atc gca tta atc aaa gag ttc gaa ggt 1907
 Met Lys Ile Ser Ser Arg Gly Ile Ala Leu Ile Lys Glu Phe Glu Gly
 1 5 10 15
 ctg cgc tta cac gct tat cgc tgc gcc gct gac gtc tgg act gtc ggt 1955
 Leu Arg Leu His Ala Tyr Arg Cys Ala Ala Asp Val Trp Thr Val Gly
 20 25 30
 tat ggc cac acg gca ggg gtt aca aag ggt gac atc atc acg gtc gat 2003
 Tyr Gly His Thr Ala Gly Val Thr Lys Gly Asp Ile Ile Thr Val Asp
 35 40 45
 gaa gcc cag acg atg ctg aca aac gat att acc gta ttt gaa cgg gcg 2051
 Glu Ala Gln Thr Met Leu Thr Asn Asp Ile Thr Val Phe Glu Arg Ala
 50 55 60
 gtc agt cag gcc gtc gcg gtt cct ctg aat cag tcg caa tac gat gcc 2099
 Val Ser Gln Ala Val Ala Val Pro Leu Asn Gln Ser Gln Tyr Asp Ala
 65 70 75 80
 ctg gtt tct ttg gtt ttt aat att ggc cag ggg aat ttt aaa cgc tct 2147
 Leu Val Ser Leu Val Phe Asn Ile Gly Gln Gly Asn Phe Lys Arg Ser
 85 90 95
 acc ttg ttg aaa aaa ctc aac aaa cag gac tat gtc ggc gcc ggg aac 2195
 Thr Leu Leu Lys Lys Leu Asn Lys Gln Asp Tyr Val Gly Ala Gly Asn
 100 105 110
 gag ttt tta cgc tgg acc cgg gcc aat ggg aag gtc ctt ccc gga ctg 2243

Glu Phe Leu Arg Trp Thr Arg Ala Asn Gly Lys Val Leu Pro Gly Leu
 115 120 125
 att cgc cga cgc gaa gct gaa cgg gtg ttg ttt gag aaa ctg ggt gca 2291
 Ile Arg Arg Arg Glu Ala Glu Arg Val Leu Phe Glu Lys Leu Gly Ala
 130 135 140
 taa ccctttgcga cgtaccacaca agatgaagat aacaccgcgt actgagcggc 2344
 145
 ggcgcaacaa tgaataaatg actgtgtacg gcctgtcctt cacaacggat gggaccatca 2404
 acgtaa tga atg agg caa gac att atg tat aat att gat gat att ctg 2452
 Met Arg Gln Asp Ile Met Tyr Asn Ile Asp Asp Ile Leu
 150 155
 gag aaa gtg aat gct cca cga gca cgc ctg tca gaa gaa aac gat aca 2500
 Glu Lys Val Asn Ala Pro Arg Ala Arg Leu Ser Glu Glu Asn Asp Thr
 160 165 170 175
 gcg gtg acg ctg acg gat tta ttc tcg cgt tcg ttt ccc gag gtc aaa 2548
 Ala Val Thr Leu Thr Asp Leu Phe Ser Arg Ser Phe Pro Glu Val Lys
 180 185 190
 aaa atc act ggc gac agc ctg tca tgg gga gag gtc tgc tat ctg tac 2596
 Lys Ile Thr Gly Asp Ser Leu Ser Trp Gly Glu Val Cys Tyr Leu Tyr
 195 200 205
 agt cag gcg cag cac gaa cag aaa gaa aac cgg ctc acc gaa tcc cgt 2644
 Ser Gln Ala Gln His Glu Gln Lys Glu Asn Arg Leu Thr Glu Ser Arg
 210 215 220
 att ctg gcc cgg gcg aat ccc cta ctg gtg aat gcc gtt cgc ctg gga 2692
 Ile Leu Ala Arg Ala Asn Pro Leu Leu Val Asn Ala Val Arg Leu Gly
 225 230 235
 ata cgg cag gca gcc ggc agt cgc agc tat gat gac tgg ttt ggc tcc 2740
 Ile Arg Gln Ala Ala Gly Ser Arg Ser Tyr Asp Asp Trp Phe Gly Ser
 240 245 250 255
 cgc gca gac cgt ttc gcc cgc ccc ggc tcg gtg gcc tcc atg ttc tca 2788
 Arg Ala Asp Arg Phe Ala Arg Pro Gly Ser Val Ala Ser Met Phe Ser
 260 265 270
 ccg gcg gcg tat ctg acc gag ctg tac cgt gag gcg aag gac ctg cat 2836
 Pro Ala Ala Tyr Leu Thr Glu Leu Tyr Arg Glu Ala Lys Asp Leu His
 275 280 285
 ccg gac acc tcg ctg ttc cgg ctg gac atc cgg cgt ccc gac ctg gcg 2884
 Pro Asp Thr Ser Leu Phe Arg Leu Asp Ile Arg Arg Pro Asp Leu Ala
 290 295 300
 gcg ctg gcc ctt agc cag aat aat atg gac gac gag ctc tcc acc ctg 2932
 Ala Leu Ala Leu Ser Gln Asn Asn Met Asp Asp Glu Leu Ser Thr Leu
 305 310 315
 agc ctg tcc aat gag cta ctg tat cgc ggt atc ggg gca gcg gaa ggg 2980

gaa gga aag tta aat gat ggg atc aca ata gga ttg agt agg aaa ggg	3700
Glu Gly Lys Leu Asn Asp Gly Ile Thr Ile Gly Leu Ser Arg Lys Gly	
560 565 570 575	
ggg gga tat tac tca aca gta aac ttc act ctg att gaa tat gat cct	3748
Gly Gly Tyr Tyr Ser Thr Val Asn Phe Thr Leu Ile Glu Tyr Asp Pro	
580 585 590	
gcg ata ttc att ctt aaa tta aat aaa gtt atc cgc cta tac aag gcc	3796
Ala Ile Phe Ile Leu Lys Leu Asn Lys Val Ile Arg Leu Tyr Lys Ala	
595 600 605	
acg ggc atg acc acg gcg gaa ata tat caa atc acc aat att ctt aat	3844
Thr Gly Met Thr Thr Ala Glu Ile Tyr Gln Ile Thr Asn Ile Leu Asn	
610 615 620	
aac ggt ctc acc att gac cat gcg gtc ctg agt aaa atc ttc ctg gtc	3892
Asn Gly Leu Thr Ile Asp His Ala Val Leu Ser Lys Ile Phe Leu Val	
625 630 635	
cgt tac ctg atg cgt cac tat cag ctt gat gtg gcc cgg tca ctg ata	3940
Arg Tyr Leu Met Arg His Tyr Gln Leu Asp Val Ala Arg Ser Leu Ile	
640 645 650 655	
ttg tgc aac gga acc atc agt gac cag gcg ttc agc ggc gaa acc ggc	3988
Leu Cys Asn Gly Thr Ile Ser Asp Gln Ala Phe Ser Gly Glu Thr Gly	
660 665 670	
ctg ttc acc acg ctg ttc aac acc cca ccg ctg aac ggc cag ctg ttt	4036
Leu Phe Thr Thr Leu Phe Asn Thr Pro Pro Leu Asn Gly Gln Leu Phe	
675 680 685	
tct gca gat gat acc ccc ctc gac tta cgc tct gaa gca ccg gag gat	4084
Ser Ala Asp Asp Thr Pro Leu Asp Leu Arg Ser Glu Ala Pro Glu Asp	
690 695 700	
gct ttc cgt ctc agc gta ctg aaa cgc gca ttt aac atc agc gcc tcg	4132
Ala Phe Arg Leu Ser Val Leu Lys Arg Ala Phe Asn Ile Ser Ala Ser	
705 710 715	
ggg ctt tcc acg ctc tgg cag ttg gcc agc ggt gac agc agc gct ggg	4180
Gly Leu Ser Thr Leu Trp Gln Leu Ala Ser Gly Asp Ser Ser Ala Gly	
720 725 730 735	
ttt agc tgc tct gct gac aat atc gcc gca ctc tac cga gtg aaa ctc	4228
Phe Ser Cys Ser Ala Asp Asn Ile Ala Ala Leu Tyr Arg Val Lys Leu	
740 745 750	
ctg gct gac atc cac gac cta tcc gct ggt gag ctg tca atg ttg ctg	4276
Leu Ala Asp Ile His Asp Leu Ser Ala Gly Glu Leu Ser Met Leu Leu	
755 760 765	
tcc gtc tcc cct ttc agc ggg gtg gcc gcc ggc tcg ctg tcc gat aat	4324
Ser Val Ser Pro Phe Ser Gly Val Ala Ala Gly Ser Leu Ser Asp Asn	
770 775 780	
gag ctg acg cag ttt ctg tac cag acc acc acc tgg ctc acg gag cag	4372

Glu Leu Thr Gln Phe Leu Tyr Gln Thr Thr Thr Trp Leu Thr Glu Gln	
785 790 795	
ggc tgg acg gtc agc gat gtg ttc ctg atg ctg acg acg cag tac ggt	4420
Gly Trp Thr Val Ser Asp Val Phe Leu Met Leu Thr Thr Gln Tyr Gly	
800 805 810 815	
acc ctg ctg acc ccc gac att gag aac ctg ctc gct tcc ctg cgc aac	4468
Thr Leu Leu Thr Pro Asp Ile Glu Asn Leu Leu Ala Ser Leu Arg Asn	
820 825 830	
gga ctg tgg ggc cgt gag ctg ttc ccg gaa acg ctc ccc ggc gat ggc	4516
Gly Leu Ser Gly Arg Glu Leu Phe Pro Glu Thr Leu Pro Gly Asp Gly	
835 840 845	
gct ccc ttt att gcc gcc gcc atg cag ctg gac gcc acg gat acg gcg	4564
Ala Pro Phe Ile Ala Ala Ala Met Gln Leu Asp Ala Thr Asp Thr Ala	
850 855 860	
aag gcg atg ctg act tgg gcg gac cag ttg aag cca gag ggg ctg acg	4612
Lys Ala Met Leu Thr Trp Ala Asp Gln Leu Lys Pro Glu Gly Leu Thr	
865 870 875	
ctg acg gaa ttt att ctt ttg gtg atg aat gcc gcc cca aat gac gag	4660
Leu Thr Glu Phe Ile Leu Leu Val Met Asn Ala Ala Pro Asn Asp Glu	
880 885 890 895	
cag gcg ggc cag atg gca ggg ttc tgc caa gcc ctg tgg caa ctg gca	4708
Gln Ala Gly Gln Met Ala Gly Phe Cys Gln Ala Leu Trp Gln Leu Ala	
900 905 910	
ctg atc atc cgc agc acc ggc ctc agc acg cgc gag ctg acg ctg ctg	4756
Leu Ile Ile Arg Ser Thr Gly Leu Ser Thr Arg Glu Leu Thr Leu Leu	
915 920 925	
gtc agc cag ccg gga cgc ttc cgc aca gga tgg cac cat ctg ccc cat	4804
Val Ser Gln Pro Gly Arg Phe Arg Thr Gly Trp His His Leu Pro His	
930 935 940	
gac ctg ccg gcg ctt cgc gac att acg cgt ttt cat gcc gtc gtt aac	4852
Asp Leu Pro Ala Leu Arg Asp Ile Thr Arg Phe His Ala Val Val Asn	
945 950 955	
cgc agc ggc agc cat gcc ggg gag gtc ctg acc gca ctt gag acc gga	4900
Arg Ser Gly Ser His Ala Gly Glu Val Leu Thr Ala Leu Glu Thr Gly	
960 965 970 975	
gaa ctg tgg tca gcc ctg ctg gcc ccg gcc ctg tca cag aat gag cag	4948
Glu Leu Ser Ser Ala Leu Leu Ala Arg Ala Leu Ser Gln Asn Glu Gln	
980 985 990	
gat gtg acc ggc gcc ttg gcg cag gtg agg ggg gcc ggt gaa cag gac	4996
Asp Val Thr Gly Ala Leu Ala Gln Val Arg Gly Ala Gly Glu Gln Asp	
995 1000 1005	
aac agc gtg ttc acc tcc tgg gaa gag gtg gac cag gct gag cag tgg	5044
Asn Ser Val Phe Thr Ser Trp Glu Glu Val Asp Gln Ala Glu Gln Trp	
1010 1015 1020	

ctg gac atg agt gag acc ctg tcc att acg cca tcc ggt ctg gct agc 5092
 Leu Asp Met Ser Glu Thr Leu Ser Ile Thr Pro Ser Gly Leu Ala Ser
 1025 1030 1035

ctg att gcc ctg aag tac atc aat gtg tcc gat gac agt gca ccg ttg 5140
 Leu Ile Ala Leu Lys Tyr Ile Asn Val Ser Asp Asp Ser Ala Pro Leu
 1040 1045 1050 1055

tac agc cag tgg cag gtg gta tcc ggt ctg ctg cag gcc ggg ctg aaa 5188
 Tyr Ser Gln Trp Gln Val Val Ser Gly Leu Leu Gln Ala Gly Leu Lys
 1060 1065 1070

agc agc cag agc tcg gcg ctg cac gat tat ctg gag gag ggg acc agc 5236
 Ser Ser Gln Ser Ser Ala Leu His Asp Tyr Leu Glu Glu Gly Thr Ser
 1075 1080 1085

agc gcc ctt tgt gcg tat tat ctg cgt aat ctg gca ccg aac atg gta 5284
 Ser Ala Leu Cys Ala Tyr Tyr Leu Arg Asn Leu Ala Pro Asn Met Val
 1090 1095 1100

tcc ggg cgc gat gac ctc ttc ggg tat ctg ctg ctg gat aat cag gtg 5332
 Ser Gly Arg Asp Asp Leu Phe Gly Tyr Leu Leu Leu Asp Asn Gln Val
 1105 1110 1115

tca gcc aag gta aaa acc acc cgc att gcg gag gcc atc gcc ggc ata 5380
 Ser Ala Lys Val Lys Thr Thr Arg Ile Ala Glu Ala Ile Ala Gly Ile
 1120 1125 1130 1135

cgg ctg tat atc aac cgg gcc ctt aac gga ata gaa ctc agc gcc atg 5428
 Arg Leu Tyr Ile Asn Arg Ala Leu Asn Gly Ile Glu Leu Ser Ala Met
 1140 1145 1150

gca gag gtg agg ggg cgt cag ttt ttc act gac tgg gat acg ttc aac 5476
 Ala Glu Val Arg Gly Arg Gln Phe Phe Thr Asp Trp Asp Thr Phe Asn
 1155 1160 1165

aaa cgt tac agc acc tgg gcg ggc gtc tca gag ctg gtt tac tat ccg 5524
 Lys Arg Tyr Ser Thr Trp Ala Gly Val Ser Glu Leu Val Tyr Tyr Pro
 1170 1175 1180

gaa aac tac ctc gac ccg acg gtc cgt atc ggg cag acc ggc atg atg 5572
 Glu Asn Tyr Leu Asp Pro Thr Val Arg Ile Gly Gln Thr Gly Met Met
 1185 1190 1195

gac acc ctg ctg cag tct gtc agc cag agc agt atc aac cgc gat acc 5620
 Asp Thr Leu Leu Gln Ser Val Ser Gln Ser Ser Ile Asn Arg Asp Thr
 1200 1205 1210 1215

gtg gag gat gcc ttt aaa acc tat ctg acc acg ttt gag cag att gcc 5668
 Val Glu Asp Ala Phe Lys Thr Tyr Leu Thr Thr Phe Glu Gln Ile Ala
 1220 1225 1230

aat ctg aac act gtc agc gga tat cac gat aac gcc agc atg acg cag 5716
 Asn Leu Asn Thr Val Ser Gly Tyr His Asp Asn Ala Ser Met Thr Gln
 1235 1240 1245

ggg act aca tgg tat gtg ggt cgc agc atc aca gat cag act aac tgg 5764

Gly Thr Thr Trp Tyr Val Gly Arg Ser Ile Thr Asp Gln Thr Asn Trp	
1250 1255 1260	
tac tgg cgc agc gcc aac cac agc aaa atc caa gac tca atg atg ccc	5812
Tyr Trp Arg Ser Ala Asn His Ser Lys Ile Gln Asp Ser Met Met Pro	
1265 1270 1275	
gcg aat gcc tgg acc gga tgg aca aaa att aac tgc gga atg aat ccg	5860
Ala Asn Ala Trp Thr Gly Trp Thr Lys Ile Asn Cys Gly Met Asn Pro	
1280 1285 1290 1295	
tgg tca gat ctt gtg tgc tgc gtg ttt ttc aac agt cgc ctt tat gtc	5908
Trp Ser Asp Leu Val Cys Ser Val Phe Phe Asn Ser Arg Leu Tyr Val	
1300 1305 1310	
gtc tgg gtc gaa gag aat cag tct gct gat acg gag gca gag agc acg	5956
Val Trp Val Glu Glu Asn Gln Ser Ala Asp Thr Glu Ala Glu Ser Thr	
1315 1320 1325	
aca acc acg cag cag agc tac acg ctg aaa ctg tgc ttc cgg cgc tac	6004
Thr Thr Thr Gln Gln Ser Tyr Thr Leu Lys Leu Ser Phe Arg Arg Tyr	
1330 1335 1340	
gac ggt aca tgg agt tcc ccg gtg tgc ttc gac att acc ggc aac atc	6052
Asp Gly Thr Trp Ser Ser Pro Val Ser Phe Asp Ile Thr Gly Asn Ile	
1345 1350 1355	
gca ttt ccg gaa acg cag gcc atg cat gtg acc tgt aat ccc ctg act	6100
Ala Phe Pro Glu Thr Gln Gly Met His Val Thr Cys Asn Pro Leu Thr	
1360 1365 1370 1375	
gag cag ctc tat tgc gcg ttt tac tcc gtc acc agc aag ccg gac ttt	6148
Glu Gln Leu Tyr Cys Ala Phe Tyr Ser Val Thr Ser Lys Pro Asp Phe	
1380 1385 1390	
gat aac gct cag ctg att tct gtg gat aat gat atg acg cta aat gtc	6196
Asp Asn Ala Gln Leu Ile Ser Val Asp Asn Asp Met Thr Leu Asn Val	
1395 1400 1405	
atc tca gat ata ggg att ttt aag agc gtc agt cac gaa ttt aat acg	6244
Ile Ser Asp Ile Gly Ile Phe Lys Ser Val Ser His Glu Phe Asn Thr	
1410 1415 1420	
agc act gag aaa ttt att aat aat gtt ttt tca gac cct tcc gct aat	6292
Ser Thr Glu Lys Phe Ile Asn Asn Val Phe Ser Asp Pro Ser Ala Asn	
1425 1430 1435	
tat ttt gtc agt gca acg agt tta att gat gat gtt atc cac agc gat	6340
Tyr Phe Val Ser Ala Thr Ser Leu Ile Asp Asp Val Ile His Ser Asp	
1440 1445 1450 1455	
ttc tca ctc ctt aat tct aaa act aca agt act gtt ttt act aat gaa	6388
Phe Ser Leu Leu Asn Ser Lys Thr Thr Ser Thr Val Phe Thr Asn Glu	
1460 1465 1470	
gat tcc tct ctt ttg acg cca gag ctt cat att aca gca aat gtt tgc	6436
Asp Ser Ser Leu Leu Thr Pro Glu Leu His Ile Thr Ala Asn Val Ser	
1475 1480 1485	

tgt ttt gtt agt act gct ggc atc gcc act caa tct acc ata gaa aaa Cys Phe Val Ser Thr Ala Gly Ile Ala Thr Gln Ser Thr Ile Glu Lys 1490 1495 1500	6484
ttc gtt cag gca ggg ata gaa ttt gag gaa att aat ttt tat gca ggc Phe Val Gln Ala Gly Ile Glu Phe Glu Glu Ile Asn Phe Tyr Ala Gly 1505 1510 1515	6532
cag gcc gcc ggc gga ttt gac gga ttt gtg gga gtg gat gtt tct aat Gln Ala Ala Gly Gly Phe Asp Gly Phe Val Gly Val Asp Val Ser Asn 1520 1525 1530 1535	6580
tca aaa gta tac cag gtc gga aaa gaa gca gtt ggt gtc act gta aaa Ser Lys Val Tyr Gln Val Gly Lys Glu Ala Val Gly Val Thr Val Lys 1540 1545 1550	6628
tct tat tcc gtc act ggc gtt agt ggt tct gtt gag tta ttt att gat Ser Tyr Ser Val Thr Gly Val Ser Gly Ser Val Glu Leu Phe Ile Asp 1555 1560 1565	6676
tca tca aat aaa tac ttc agc gga att ttg tca gat aaa atg ata acc Ser Ser Asn Lys Tyr Phe Ser Gly Ile Leu Ser Asp Lys Met Ile Thr 1570 1575 1580	6724
gct tta att agc ggc agt aca tca aaa gtt aat tac gtg tcg tct att Ala Leu Ile Ser Gly Ser Thr Ser Lys Val Asn Tyr Val Ser Ser Ile 1585 1590 1595	6772
ggc tct caa gat ttt tgg agt gta aag tcg ctc atg ccg gca ctt cag Gly Ser Gln Asp Phe Trp Ser Val Lys Ser Leu Met Pro Ala Leu Gln 1600 1605 1610 1615	6820
ata tat gaa tta atc gat gat atc ata ctg aca tcc ggc gta aat ggg Ile Tyr Glu Leu Ile Asp Asp Ile Ile Leu Thr Ser Gly Val Asn Gly 1620 1625 1630	6868
act gaa att aaa tcc tgg cct tcc gct gaa tgg tat aat gat aag ctg Thr Glu Ile Lys Ser Trp Pro Ser Ala Glu Trp Tyr Asn Asp Lys Leu 1635 1640 1645	6916
agt ctg caa tcc ggg aat aat ctt ttc aac acc aaa tcg ctg agt ttt Ser Leu Gln Ser Gly Asn Asn Leu Phe Asn Thr Lys Ser Leu Ser Phe 1650 1655 1660	6964
acc gtt aat acc agt gat att gtt gaa gat gag ttt gac gtg acg ttt Thr Val Asn Thr Ser Asp Ile Val Glu Asp Glu Phe Asp Val Thr Phe 1665 1670 1675	7012
acg ttc acc gct gtc gat cag aat aac gtc gtg ctg gcc gcc cgg acg Thr Phe Thr Ala Val Asp Gln Asn Asn Val Val Leu Ala Ala Arg Thr 1680 1685 1690 1695	7060
gcc ata tta acc gtc att cga aac att aat aat gac act tcc gtt atc Ala Ile Leu Thr Val Ile Arg Asn Ile Asn Asn Asp Thr Ser Val Ile 1700 1705 1710	7108
gca tta cgt aaa aat acg cgt ggc gcg cag tat att cgt ttc act gcg	7156

Ala Leu Arg Lys Asn Thr Arg Gly Ala Gln Tyr Ile Arg Phe Thr Ala	
1715 1720 1725	
ggt aac gat gtg gcg ctt att cgc ctc aac acc ctc ttt gcc cgc caa	7204
Gly Asn Asp Val Ala Leu Ile Arg Leu Asn Thr Leu Phe Ala Arg Gln	
1730 1735 1740	
ctg gtc gac cgg gcg aat acc ggg att gac acc att ctt tcc atg gag	7252
Leu Val Asp Arg Ala Asn Thr Gly Ile Asp Thr Ile Leu Ser Met Glu	
1745 1750 1755	
acc cag agg ctt acc gaa ccc gcc ctg gaa gag ggg agt gat gtg ttt	7300
Thr Gln Arg Leu Thr Glu Pro Ala Leu Glu Gly Ser Asp Val Phe	
1760 1765 1770 1775	
atg gac ttc tcc gga gcc aat gcc ctc tat ttc tgg gag ctg ttc tat	7348
Met Asp Phe Ser Gly Ala Asn Ala Leu Tyr Phe Trp Glu Leu Phe Tyr	
1780 1785 1790	
tac acg ccg atg atg gtg ttc cag cgg ttg ttg cag gaa cag cac ttc	7396
Tyr Thr Pro Met Met Val Phe Gln Arg Leu Leu Gln Glu Gln His Phe	
1795 1800 1805	
ccg gaa gcc acc cgc tgg ctg cag tat gtc tgg aac ccg gcc ggg cac	7444
Pro Glu Ala Thr Arg Trp Leu Gln Tyr Val Trp Asn Pro Ala Gly His	
1810 1815 1820	
gtg gta aac ggg gtg ctg cag aat tac acc tgg aat gtc cgt ccg ctg	7492
Val Val Asn Gly Val Leu Gln Asn Tyr Thr Trp Asn Val Arg Pro Leu	
1825 1830 1835	
gag gag gac acc ggc tgg aac gac tgc ccg ctg gac tcc att gac ccc	7540
Glu Glu Asp Thr Gly Trp Asn Asp Ser Pro Leu Asp Ser Ile Asp Pro	
1840 1845 1850 1855	
gat gca ata gcc cag tac gac ccc atg cat tac aag gtc gcc acc ttt	7588
Asp Ala Ile Ala Gln Tyr Asp Pro Met His Tyr Lys Val Ala Thr Phe	
1860 1865 1870	
atg tgc tac ctc gac ctg ctg att gcc cgc ggt gat gcc gcc tac cgg	7636
Met Ser Tyr Leu Asp Leu Leu Ile Ala Arg Gly Asp Ala Ala Tyr Arg	
1875 1880 1885	
ctg ctc gag cgg gac acc ctt aac gag gcc cgg atg tgg tac gtc cag	7684
Leu Leu Glu Arg Asp Thr Leu Asn Glu Ala Arg Met Trp Tyr Val Gln	
1890 1895 1900	
gcc ctg aac ctt ctg ggc gac gag ccc tat att tcc ttt gac gcc gac	7732
Ala Leu Asn Leu Leu Gly Asp Glu Pro Tyr Ile Ser Phe Asp Ala Asp	
1905 1910 1915	
tgg tgc gcg ttg acc ctg ggt gac gca gcc agc gag gtg acg cga cgc	7780
Trp Ser Ala Leu Thr Leu Gly Asp Ala Ala Ser Glu Val Thr Arg Arg	
1920 1925 1930 1935	
gat tac cag gag gcc ctg ctg gcc gtg cgc cgg ttg gtg ccc gct ccc	7828
Asp Tyr Gln Glu Ala Leu Leu Ala Val Arg Arg Leu Val Pro Ala Pro	
1940 1945 1950	

gag aca cgg acg gcg aat tcc ctg acg gca ctg ttc ctc ccg cag cag Glu Thr Arg Thr Ala Asn Ser Leu Thr Ala Leu Phe Leu Pro Gln Gln 1955 1960 1965	7876
aac gag gtg ctc aaa ggc tac tgg caa acc ttg gca cag cgg ctc cat Asn Glu Val Leu Lys Gly Tyr Trp Gln Thr Leu Ala Gln Arg Leu His 1970 1975 1980	7924
aac ctg cgc cac aac ctc tcc att gac ggc cag ccg ctt tcc ctg tcc Asn Leu Arg His Asn Leu Ser Ile Asp Gly Gln Pro Leu Ser Leu Ser 1985 1990 1995	7972
gtc tac gcc acg ccg tcc gaa ccg tcc gcc ctg cag agt gcc gtc gtc Val Tyr Ala Thr Pro Ser Glu Pro Ser Ala Leu Gln Ser Ala Val Val 2000 2005 2010 2015	8020
aac agc gcg cag ggt gct gca gca ctg ccg gcc gcg gtg atg ccg ctt Asn Ser Ala Gln Gly Ala Ala Ala Leu Pro Ala Ala Val Met Pro Leu 2020 2025 2030	8068
tac agt ttc ccg gtc atg ctg gag aac gcc ccg ggg atg gtg agc ctg Tyr Ser Phe Pro Val Met Leu Glu Asn Ala Arg Gly Met Val Ser Leu 2035 2040 2045	8116
ctg acc ggg ttc ggc aac aca ctg ctc ggt att acc gag cgt cag gat Leu Thr Gly Phe Gly Asn Thr Leu Leu Gly Ile Thr Glu Arg Gln Asp 2050 2055 2060	8164
gcg gag gcg ctg gcc aaa ctg ctg cag acc cag ggc agt gaa ctg ata Ala Glu Ala Leu Ala Lys Leu Leu Gln Thr Gln Gly Ser Glu Leu Ile 2065 2070 2075	8212
cgc cag ggc ctt cgc cag cag gat aac gtc ctc gag gaa atc gat gcg Arg Gln Gly Leu Arg Gln Gln Asp Asn Val Leu Glu Glu Ile Asp Ala 2080 2085 2090 2095	8260
gat att gcc gcc ctg gag gag agc cgc cgc ggc gcg cag atg cgt ttt Asp Ile Ala Ala Leu Glu Glu Ser Arg Arg Gly Ala Gln Met Arg Phe 2100 2105 2110	8308
gaa cgt tac aaa gtg ttg tac gag gcg gac gtc aac acc ggc gaa aaa Glu Arg Tyr Lys Val Leu Tyr Glu Ala Asp Val Asn Thr Gly Glu Lys 2115 2120 2125	8356
cag gcc atg gac ttg tac ctc agt tcc tcc gtg ctg tcc gca tca acc Gln Ala Met Asp Leu Tyr Leu Ser Ser Ser Val Leu Ser Ala Ser Thr 2130 2135 2140	8404
gcc gcg ctc ttt ttg gcc gag gcc gcg gcc gat atg ctg ccc aat att Ala Ala Leu Phe Leu Ala Glu Ala Ala Ala Asp Met Leu Pro Asn Ile 2145 2150 2155	8452
tac ggg ctg gcc gtc ggg ggc tcc cgc tat ggg gca cta ttt aaa gcc Tyr Gly Leu Ala Val Gly Gly Ser Arg Tyr Gly Ala Leu Phe Lys Ala 2160 2165 2170 2175	8500
acc gcc atc ggc atc cag gtg tcc tcc gat gcc acc cgc ata tca gcg	8548

Thr Ala Ile Gly Ile Gln Val Ser Ser Asp Ala Thr Arg Ile Ser Ala	
2180 2185 2190	
gac aaa atc agc cag tcg gaa gtg tac cgc cgt cgc cgg gag gag tgg	8596
Asp Lys Ile Ser Gln Ser Glu Val Tyr Arg Arg Arg Arg Glu Glu Trp	
2195 2200 2205	
gaa atc cag cgt gat agt gcg cag tct gac gtg gcg cag att gat gcc	8644
Glu Ile Gln Arg Asp Ser Ala Gln Ser Asp Val Ala Gln Ile Asp Ala	
2210 2215 2220	
cag ctg gcg gcc atg gca gtg cgc cgg gaa ggg gct gag ctg cag aaa	8692
Gln Leu Ala Ala Met Ala Val Arg Arg Glu Gly Ala Glu Leu Gln Lys	
2225 2230 2235	
act tac ctt gag acc cag cag acc cag gca cag gcg cag ttg gca ttc	8740
Thr Tyr Leu Glu Thr Gln Gln Thr Gln Ala Gln Ala Gln Leu Ala Phe	
2240 2245 2250 2255	
ctg cag agt aag ttc aac aat acg gct ctg tac agc tgg ctg cgg ggc	8788
Leu Gln Ser Lys Phe Asn Asn Thr Ala Leu Tyr Ser Trp Leu Arg Gly	
2260 2265 2270	
agg ttg tcc gcc att tat tac cag ttc tat gac ctg gca gta tcc cgc	8836
Arg Leu Ser Ala Ile Tyr Tyr Gln Phe Tyr Asp Leu Ala Val Ser Arg	
2275 2280 2285	
tgc ctg atg gcg caa cag gcc tgg cag tgg gat aaa ttc gag act agg	8884
Cys Leu Met Ala Gln Gln Ala Trp Gln Trp Asp Lys Phe Glu Thr Arg	
2290 2295 2300	
tcg ttt atc cag ccg ggg gcc tgg atg ggg gca aat gcc ggt ctg ctg	8932
Ser Phe Ile Gln Pro Gly Ala Trp Met Gly Ala Asn Ala Gly Leu Leu	
2305 2310 2315	
gcc ggg gaa acc ctg atg ctg aat ctg gcg cag atg gag cag gcc tgg	8980
Ala Gly Glu Thr Leu Met Leu Asn Leu Ala Gln Met Glu Gln Ala Trp	
2320 2325 2330 2335	
ctg acg ggg gat gag cgg gca ata gag gtg acg cgg acg gtc tgc ctg	9028
Leu Thr Gly Asp Glu Arg Ala Ile Glu Val Thr Arg Thr Val Cys Leu	
2340 2345 2350	
tcg gag gtc tat acc agc ctc gcg gag gat gcg gca ttc tct ctg gcc	9076
Ser Glu Val Tyr Thr Ser Leu Ala Glu Asp Ala Ala Phe Ser Leu Ala	
2355 2360 2365	
gac aag gtg gtg gaa ctg gtc agt aac ggt tcg ggc agt gcg ggt acg	9124
Asp Lys Val Val Glu Leu Val Ser Asn Gly Ser Gly Ser Ala Gly Thr	
2370 2375 2380	
aaa agc aac gga tta cag atg gat caa cag caa ctc gag gcc acc ctg	9172
Lys Ser Asn Gly Leu Gln Met Asp Gln Gln Gln Leu Glu Ala Thr Leu	
2385 2390 2395	
aaa ctg gct gac ctc ggt atc ggc aac gat tac ccg gtc tcc ctt ggc	9220
Lys Leu Ala Asp Leu Gly Ile Gly Asn Asp Tyr Pro Val Ser Leu Gly	
2400 2405 2410 2415	

acc atg agg cgc atc aaa caa ata agc gtc acg ctc ccg gcg ctg gtc Thr Met Arg Arg Ile Lys Gln Ile Ser Val Thr Leu Pro Ala Leu Val 2420 2425 2430	9268
ggc ccc tat cag gac gtc cgt gcg gtt ctc agc tac ggc gga agt atg Gly Pro Tyr Gln Asp Val Arg Ala Val Leu Ser Tyr Gly Gly Ser Met 2435 2440 2445	9316
gtc atg ccc cgg ggt tgc agc gcg ctg gcg gtc tca cac gga atg aac Val Met Pro Arg Gly Cys Ser Ala Leu Ala Val Ser His Gly Met Asn 2450 2455 2460	9364
gac agc ggc caa ttc caa ctg gat ttc aat gac ccg cgt tac ctg ccg Asp Ser Gly Gln Phe Gln Leu Asp Phe Asn Asp Pro Arg Tyr Leu Pro 2465 2470 2475	9412
ttt gaa gga ctt cca gtt gat gac aca ggg acc ctg aca ctg agc ttc Phe Glu Gly Leu Pro Val Asp Asp Thr Gly Thr Leu Thr Leu Ser Phe 2480 2485 2490 2495	9460
ccg gat gct gac ggc aaa caa cag gcg atg ctc ctc agt ctg agc gac Pro Asp Ala Asp Gly Lys Gln Gln Ala Met Leu Leu Ser Leu Ser Asp 2500 2505 2510	9508
atc atc ctg cat atc cgt tac acc att atc agc tga tag gtatcaacat Ile Ile Leu His Ile Arg Tyr Thr Ile Ile Ser 2515 2520	9557
agegcaggcc cccgaacgag ggccctgcgag gagactgagc atg caa aat cat caa Met Gln Asn His Gln 2525	9612
gac atg gcc att act gcc ccc acg ttg cct tcc ggg ggc ggt gcg gtc Asp Met Ala Ile Thr Ala Pro Thr Leu Pro Ser Gly Gly Gly Ala Val 2530 2535 2540 2545	9660
acc ggg ctc aag ggt gat atc gcg gcg gca ggg ccg gat ggt gcg gcg Thr Gly Leu Lys Gly Asp Ile Ala Ala Ala Gly Pro Asp Gly Ala Ala 2550 2555 2560	9708
acc ctg agt att ccc ttg ccg gtt agc ccc ggt ccg ggt tac gcc ccc Thr Leu Ser Ile Pro Leu Pro Val Ser Pro Gly Arg Gly Tyr Ala Pro 2565 2570 2575	9756
act ggg gca ctt aat tat cac agc cgg tcg ggg aac ggc ccc ttt ggc Thr Gly Ala Leu Asn Tyr His Ser Arg Ser Gly Asn Gly Pro Phe Gly 2580 2585 2590	9804
att ggc tgg ggt atc ggc ggt gct gct gtc cag cgt cgt acg cgc aac Ile Gly Trp Gly Ile Gly Gly Ala Ala Val Gln Arg Arg Thr Arg Asn 2595 2600 2605	9852
gga gca cct acc tac gat gat act gat gaa ttc acc ggt ccg gac ggt Gly Ala Pro Thr Tyr Asp Asp Thr Asp Glu Phe Thr Gly Pro Asp Gly 2610 2615 2620 2625	9900
gag gtg ctg gtg ccg gca ctc acg gct gct ggc acc caa gaa gca cgg	9948

Glu Val Leu Val Pro Ala Leu Thr Ala Ala Gly Thr Gln Glu Ala Arg	
2630 2635 2640	
cag gcc acc tca cta ctg ggg ata aac cca ggc gga agc ttc aac gtt	9996
Gln Ala Thr Ser Leu Leu Gly Ile Asn Pro Gly Gly Ser Phe Asn Val	
2645 2650 2655	
cag gtt tac cgt tca cgt acg gag ggt agt ctc agc cgc ctt gag cgt	10044
Gln Val Tyr Arg Ser Arg Thr Glu Gly Ser Leu Ser Arg Leu Glu Arg	
2660 2665 2670	
tgg ctg ccc gcc gac gag aca gaa acg gaa ttt tgg gtg tta tat acc	10092
Trp Leu Pro Ala Asp Glu Thr Glu Thr Glu Phe Trp Val Leu Tyr Thr	
2675 2680 2685	
cct gac gga cag gtg gct ctg ctg ggc cga aat gcg cag gct cgc atc	10140
Pro Asp Gly Gln Val Ala Leu Leu Gly Arg Asn Ala Gln Ala Arg Ile	
2690 2695 2700 2705	
agc aac ccc aca gcc cca aca cag acg gcg gtt tgg ctg atg gag tcc	10188
Ser Asn Pro Thr Ala Pro Thr Gln Thr Ala Val Trp Leu Met Glu Ser	
2710 2715 2720	
tcg gta tca ctt acc ggc gaa cag atg tat tac caa tac cgt gcg gaa	10236
Ser Val Ser Leu Thr Gly Glu Gln Met Tyr Tyr Gln Tyr Arg Ala Glu	
2725 2730 2735	
gat gat gac ggt tgt gac gag gcg gag cgc gac gcg cac ccg cag gcc	10284
Asp Asp Asp Gly Cys Asp Glu Ala Glu Arg Asp Ala His Pro Gln Ala	
2740 2745 2750	
ggc gcc caa cgt tat ccg gtg gcg gtc tgg tat ggt aac cgt cag gcg	10332
Gly Ala Gln Arg Tyr Pro Val Ala Val Trp Tyr Gly Asn Arg Gln Ala	
2755 2760 2765	
gct cgg acg cta ccg gcg ctg gtg tcg aca cca tca atg gat agc tgg	10380
Ala Arg Thr Leu Pro Ala Leu Val Ser Thr Pro Ser Met Asp Ser Trp	
2770 2775 2780 2785	
ctg ttt atc ctg gtg ttt gat tat ggt gag cgt agc tcg gtg ctg tct	10428
Leu Phe Ile Leu Val Phe Asp Tyr Gly Glu Arg Ser Ser Val Leu Ser	
2790 2795 2800	
gaa gcg ccg gcc tgg caa aca cca gga agt ggg gag tgg ctg tgt cgt	10476
Glu Ala Pro Ala Trp Gln Thr Pro Gly Ser Gly Glu Trp Leu Cys Arg	
2805 2810 2815	
cag gat tgt ttt tcc ggg tat gag ttt ggt ttt aac ctg cgg act cgc	10524
Gln Asp Cys Phe Ser Gly Tyr Glu Phe Gly Phe Asn Leu Arg Thr Arg	
2820 2825 2830	
cgc ctg tgc cgt cag gtt ttg atg ttc cat tac cta ggt gtt ctg gcg	10572
Arg Leu Cys Arg Gln Val Leu Met Phe His Tyr Leu Gly Val Leu Ala	
2835 2840 2845	
ggg agt tcg gga gcg aat gat gcg cca gca ttg att tct cgc ctg ttg	10620
Gly Ser Ser Gly Ala Asn Asp Ala Pro Ala Leu Ile Ser Arg Leu Leu	
2850 2855 2860 2865	

ctg gac tac agg gaa agt cct tca ctc agt ctg ctc gag aac gtg cac Leu Asp Tyr Arg Glu Ser Pro Ser Leu Ser Leu Leu Glu Asn Val His 2870 2875 2880	10668
cag gtg gct tat gag tcg gac ggg acg tct tgt gcc ttg ccg gca ctg Gln Val Ala Tyr Glu Ser Asp Gly Thr Ser Cys Ala Leu Pro Ala Leu 2885 2890 2895	10716
gca ttg ggg tgg caa acc ttt acc ccg ccg aca ttg tcg gca tgg cag Ala Leu Gly Trp Gln Thr Phe Thr Pro Pro Thr Leu Ser Ala Trp Gln 2900 2905 2910	10764
acg cgt gac gat atg ggc aag ttg agt ttg ctt caa ccc tat cag ctt Thr Arg Asp Asp Met Gly Lys Leu Ser Leu Leu Gln Pro Tyr Gln Leu 2915 2920 2925	10812
gta gac ctt aac ggc gaa ggt gtg gtg ggt atc ctg tat cag gac agc Val Asp Leu Asn Gly Glu Gly Val Val Gly Ile Leu Tyr Gln Asp Ser 2930 2935 2940 2945	10860
ggt gcc tgg tgg tac cgt gaa ccg gta cgc cag tcg ggg gat gat ccg Gly Ala Trp Trp Tyr Arg Glu Pro Val Arg Gln Ser Gly Asp Asp Pro 2950 2955 2960	10908
gat gct gtg acc tgg ggg gcg gct gcg gcc ctg ccg aca atg ccc gct Asp Ala Val Thr Trp Gly Ala Ala Ala Leu Pro Thr Met Pro Ala 2965 2970 2975	10956
ttg cat aac agc ggc atc ctg gcg gat ctt aat ggg gat ggt cgg ctg Leu His Asn Ser Gly Ile Leu Ala Asp Leu Asn Gly Asp Gly Arg Leu 2980 2985 2990	11004
gag tgg gtc gtt acc gcc ccc ggt gtg gcg ggg atg tat gat cgc acc Glu Trp Val Val Thr Ala Pro Gly Val Ala Gly Met Tyr Asp Arg Thr 2995 3000 3005	11052
ccc ggc cgc gac tgg ttg cat ttc acc ccc ctg tca gcc ttg ccc gta Pro Gly Arg Asp Trp Leu His Phe Thr Pro Leu Ser Ala Leu Pro Val 3010 3015 3020 3025	11100
gaa tat gcg cat cca aaa gca gtg ctc gcc gat atc ctg ggg gct ggg Glu Tyr Ala His Pro Lys Ala Val Leu Ala Asp Ile Leu Gly Ala Gly 3030 3035 3040	11148
tta acg gac atg gtg ctt atc ggg ccg cgc agt gtt cgc ctc tat tcc Leu Thr Asp Met Val Leu Ile Gly Pro Arg Ser Val Arg Leu Tyr Ser 3045 3050 3055	11196
ggc aaa aac gat ggt tgg aat aaa ggg gag acc gtg cag caa acg gaa Gly Lys Asn Asp Gly Trp Asn Lys Gly Glu Thr Val Gln Gln Thr Glu 3060 3065 3070	11244
aga ctc act ctg ccg gtc ccg ggg gtt gac cca cgt acc ctc gtg gcg Arg Leu Thr Leu Pro Val Pro Gly Val Asp Pro Arg Thr Leu Val Ala 3075 3080 3085	11292
ttc agt gat atg gct ggc agt gga cag cag cat ttg acg gag gtg cgt	11340

Phe Ser Asp Met Ala Gly Ser Gly Gln Gln His Leu Thr Glu Val Arg
 3090 3095 3100 3105
 gct aat gga gta cgt tac tgg cca aac ctg ggg cac ggt cgt ttc ggt 11388
 Ala Asn Gly Val Arg Tyr Trp Pro Asn Leu Gly His Gly Arg Phe Gly
 3110 3115 3120
 cag ccg gtg aat att ccc ggt ttt agc cag tca gtg act acg ttt aac 11436
 Gln Pro Val Asn Ile Pro Gly Phe Ser Gln Ser Val Thr Thr Phe Asn
 3125 3130 3135
 cct gac cag ata ttg ctg gcc gat acc gac ggt tcc ggt acc acg gac 11484
 Pro Asp Gln Ile Leu Leu Ala Asp Thr Asp Gly Ser Gly Thr Thr Asp
 3140 3145 3150
 ctg att tat gcg atg agt gac cgg tta gtc att tat ttc aac cag agt 11532
 Leu Ile Tyr Ala Met Ser Asp Arg Leu Val Ile Tyr Phe Asn Gln Ser
 3155 3160 3165
 ggt aat tat ttc gcc gag ccg cat acg ctg ctc ttg ccg aaa ggt gtg 11580
 Gly Asn Tyr Phe Ala Glu Pro His Thr Leu Leu Leu Pro Lys Gly Val
 3170 3175 3180 3185
 cgc tat gat cgc acc tgc agt ctg caa gtg gcg gat atc cag ggg ctg 11628
 Arg Tyr Asp Arg Thr Cys Ser Leu Gln Val Ala Asp Ile Gln Gly Leu
 3190 3195 3200
 ggg gtg cct agc ctg tta ctg acg gtc ccc cat gtc gcg cct cat cac 11676
 Gly Val Pro Ser Leu Leu Leu Thr Val Pro His Val Ala Pro His His
 3205 3210 3215
 tgg gtg tgc cat tta tgc gca gac aaa ccc tgg ttg ttg aat ggc atg 11724
 Trp Val Cys His Leu Ser Ala Asp Lys Pro Trp Leu Leu Asn Gly Met
 3220 3225 3230
 aac aac aat atg ggg gcc cgg cat gca ctg cac tat cgc agt tgc gtg 11772
 Asn Asn Asn Met Gly Ala Arg His Ala Leu His Tyr Arg Ser Ser Val
 3235 3240 3245
 cag ttc tgg ctg gat gag aaa gcc gag gca ctg gcg gca ggc agt tcc 11820
 Gln Phe Trp Leu Asp Glu Lys Ala Glu Ala Leu Ala Ala Gly Ser Ser
 3250 3255 3260 3265
 cct gcc tgc tac ctg cca ttt aca ttg cat acc ctg tgg cgt tgc gtg 11868
 Pro Ala Cys Tyr Leu Pro Phe Thr Leu His Thr Leu Trp Arg Ser Val
 3270 3275 3280
 gtg cag gat gag atc acc ggt aac cgt ctg gtc agc gac gtg ctt tat 11916
 Val Gln Asp Glu Ile Thr Gly Asn Arg Leu Val Ser Asp Val Leu Tyr
 3285 3290 3295
 cgc cac ggc gtc tgg gac ggg cag gaa cgc gag ttt cgg ggg ttt ggt 11964
 Arg His Gly Val Trp Asp Gly Gln Glu Arg Glu Phe Arg Gly Phe Gly
 3300 3305 3310
 ttt gtt gag atc agg gat acc gat acc ttg gca agc cag ggt acg gcg 12012
 Phe Val Glu Ile Arg Asp Thr Asp Thr Leu Ala Ser Gln Gly Thr Ala
 3315 3320 3325

acg gaa ctg agt atg cct tct gtg agc cgg aac tgg tat gcc acc ggg Thr Glu Leu Ser Met Pro Ser Val Ser Arg Asn Trp Tyr Ala Thr Gly 3330 3335 3340 3345	12060
gta ccg gca gta gac gag cgt ctg ccg gag acg tat tgg caa aac gat Val Pro Ala Val Asp Glu Arg Leu Pro Glu Thr Tyr Trp Gln Asn Asp 3350 3355 3360	12108
gcc gcc gct ttt gcc gat ttc gcg acc cgt ttc act gtc ggt tca gga Ala Ala Ala Phe Ala Asp Phe Ala Thr Arg Phe Thr Val Gly Ser Gly 3365 3370 3375	12156
gag gat gag cag aca tat act ccg gac gac agc aag aca ttc tgg ttg Glu Asp Glu Gln Thr Tyr Thr Pro Asp Asp Ser Lys Thr Phe Trp Leu 3380 3385 3390	12204
cag cga gcc ctg aaa ggc atc ctg ctg cgc agt gag tta tac ggt gcc Gln Arg Ala Leu Lys Gly Ile Leu Leu Arg Ser Glu Leu Tyr Gly Ala 3395 3400 3405	12252
gat ggc agc agc cag gcc gat atc cct tac agc gtc act gag tct cgc Asp Gly Ser Ser Gln Ala Asp Ile Pro Tyr Ser Val Thr Glu Ser Arg 3410 3415 3420 3425	12300
ccg cag gta cgg cta gtt gaa gcg aat gga gac tac ccg gtg gtg tgg Pro Gln Val Arg Leu Val Glu Ala Asn Gly Asp Tyr Pro Val Val Trp 3430 3435 3440	12348
ccg atg ggc gcg gaa agc cgt acg tca gtt tat gaa cgg tac cac aat Pro Met Gly Ala Glu Ser Arg Thr Ser Val Tyr Glu Arg Tyr His Asn 3445 3450 3455	12396
gat cct caa tgc caa cag cag gcg gta ctc ctc agt gat gaa tac ggt Asp Pro Gln Cys Gln Gln Ala Val Leu Leu Ser Asp Glu Tyr Gly 3460 3465 3470	12444
ttc cca ctg cgt cag gtc agt gtc aat tat cca cga cgc cct ccg tcg Phe Pro Leu Arg Gln Val Ser Val Asn Tyr Pro Arg Arg Pro Pro Ser 3475 3480 3485	12492
gcg gac aat cca tat ccg gcg tcc tta ccg gcg acg ctg ttc gcc aac Ala Asp Asn Pro Tyr Pro Ala Ser Leu Pro Ala Thr Leu Phe Ala Asn 3490 3495 3500 3505	12540
agt tat gac gag cag cag cag ata tta cgc ctg ggg ttg caa cag agc Ser Tyr Asp Glu Gln Gln Ile Leu Arg Leu Gly Leu Gln Gln Ser 3510 3515 3520	12588
agt gca cat cac ctt gtt tca ctg tct gag ggg cat tgg ttg ttg ggg Ser Ala His His Leu Val Ser Leu Ser Glu Gly His Trp Leu Leu Gly 3525 3530 3535	12636
ttg gcg gag gcg tcg ccg gac gat gta ttc acg tac tct gcg gac aac Leu Ala Glu Ala Ser Arg Asp Asp Val Phe Thr Tyr Ser Ala Asp Asn 3540 3545 3550	12684
gtg ccg gaa ggg ggt ctg acg ctg gaa cac ctg ttg gcg ccc gaa agc	12732

Val Pro Glu Gly Gly Leu Thr Leu Glu His Leu Leu Ala Pro Glu Ser	
3555 3560 3565	
ctg gtc tcg gat agt cag gtc ggt acg ctg gcg ggt cag cag caa gtc	12780
Leu Val Ser Asp Ser Gln Val Gly Thr Leu Ala Gly Gln Gln Gln Val	
3570 3575 3580 3585	
tgg tat ctg gat tca caa gac gtt gcc acc gtc gct gct ccg cca ctc	12828
Trp Tyr Leu Asp Ser Gln Asp Val Ala Thr Val Ala Ala Pro Pro Leu	
3590 3595 3600	
ccc ccc aag gta gct ttt atc gaa acg gcc gtg ctg gat gag ggt atg	12876
Pro Pro Lys Val Ala Phe Ile Glu Thr Ala Val Leu Asp Glu Gly Met	
3605 3610 3615	
gtc agt tca ctg gct gcc tac att gtg gat gaa cat ctc gag caa gcc	12924
Val Ser Ser Leu Ala Ala Tyr Ile Val Asp Glu His Leu Glu Gln Ala	
3620 3625 3630	
ggg tac cgg caa tcc gga tac ctt ttc cct cga ggc agg gaa gca gaa	12972
Gly Tyr Arg Gln Ser Gly Tyr Leu Phe Pro Arg Gly Arg Glu Ala Glu	
3635 3640 3645	
cag gca ttg tgg acc cag tgt cag gga tat gtt acc tat gcc ggc gca	13020
Gln Ala Leu Trp Thr Gln Cys Gln Gly Tyr Val Thr Tyr Ala Gly Ala	
3650 3655 3660 3665	
gag cat ttc tgg cta ccg cta tcc ttt cgg gac agt atg ttg acc ggc	13068
Glu His Phe Trp Leu Pro Leu Ser Phe Arg Asp Ser Met Leu Thr Gly	
3670 3675 3680	
cca gtt acc gtg acg cgt gac gcg tac gac tgc gtc atc acg cag tgg	13116
Pro Val Thr Val Thr Arg Asp Ala Tyr Asp Cys Val Ile Thr Gln Trp	
3685 3690 3695	
cag gat gcc gca ggg att gtc acc aca gcc gac tat gac tgg cgc ttc	13164
Gln Asp Ala Ala Gly Ile Val Thr Thr Ala Asp Tyr Asp Trp Arg Phe	
3700 3705 3710	
ctg acg ccc gtc cgg gtg acg gac ccc aat gat aat ctg cag tcc gtc	13212
Leu Thr Pro Val Arg Val Thr Asp Pro Asn Asp Asn Leu Gln Ser Val	
3715 3720 3725	
act ctg gat gct ctg ggc cgg gtg acc acc ctg cga ttc tgg ggc acg	13260
Thr Leu Asp Ala Leu Gly Arg Val Thr Thr Leu Arg Phe Trp Gly Thr	
3730 3735 3740 3745	
gag aat ggt att gcc acc ggt tac agt gat gcc acg ttg tcc gtt ccg	13308
Glu Asn Gly Ile Ala Thr Gly Tyr Ser Asp Ala Thr Leu Ser Val Pro	
3750 3755 3760	
gac ggc gca gca gcc gct ctg gcg ttg acg gcg ccc cta cca gta gca	13356
Asp Gly Ala Ala Ala Leu Ala Leu Thr Ala Pro Leu Pro Val Ala	
3765 3770 3775	
cag tgt ctg gtg tat gtc acg gac agt tgg gga gat gac gac aat gag	13404
Gln Cys Leu Val Tyr Val Thr Asp Ser Trp Gly Asp Asp Asp Asn Glu	
3780 3785 3790	

aaa atg ccc ccg cac gtg gtc gtg ctg gct acc gat cgc tat gac agt 13452
 Lys Met Pro Pro His Val Val Val Leu Ala Thr Asp Arg Tyr Asp Ser
 3795 3800 3805

gat acc gga cag cag gtc cgc caa cag gtg aca ttc agt gac ggt ttt 13500
 Asp Thr Gly Gln Gln Val Arg Gln Gln Val Thr Phe Ser Asp Gly Phe
 3810 3815 3820 3825

ggg cgt gag ttg caa tcg gca acc cgg cag gcc gag ggc aac gcc tgg 13548
 Gly Arg Glu Leu Gln Ser Ala Thr Arg Gln Ala Glu Gly Asn Ala Trp
 3830 3835 3840

caa cga gga cgc gac ggc aaa ctg gtg acg gcc agt gac gga ttg ccg 13596
 Gln Arg Gly Arg Asp Gly Lys Leu Val Thr Ala Ser Asp Gly Leu Pro
 3845 3850 3855

gtc act gta gca acg aat ttc cgc tgg gcg gtc acc ggg agg gcg gag 13644
 Val Thr Val Ala Thr Asn Phe Arg Trp Ala Val Thr Gly Arg Ala Glu
 3860 3865 3870

tat gac aat aaa ggt ctg cct gtt cgg gtt tat cag ccg tat ttt ctg 13692
 Tyr Asp Asn Lys Gly Leu Pro Val Arg Val Tyr Gln Pro Tyr Phe Leu
 3875 3880 3885

gac agt tgg caa tat gtc agt gat gac agt gcc cgc cag gac ctg tat 13740
 Asp Ser Trp Gln Tyr Val Ser Asp Asp Ser Ala Arg Gln Asp Leu Tyr
 3890 3895 3900 3905

gcc gac acg cac ttt tac gat ccg acg gca cgg gaa tgg cag gtt att 13788
 Ala Asp Thr His Phe Tyr Asp Pro Thr Ala Arg Glu Trp Gln Val Ile
 3910 3915 3920

acg gca aaa ggt gaa cgg cga cag gtg ctg tat acc ccg tgg ttt gtg 13836
 Thr Ala Lys Gly Glu Arg Arg Gln Val Leu Tyr Thr Pro Trp Phe Val
 3925 3930 3935

gtc agt gaa gac gag aat gat acc gtt ggg cta aac gac gca tcc tga 13884
 Val Ser Glu Asp Glu Asn Asp Thr Val Gly Leu Asn Asp Ala Ser
 3940 3945 3950

ctgggaagga gggggggacg gtg atg agt ccg tcg ccc ctg aca ggc gct gcc 13937
 Met Ser Pro Ser Pro Leu Thr Gly Ala Ala
 3955 3960

ctg atg gag aca aag atg aaa ata cac tat cag gtt gcg gcg gtt gtg 13985
 Leu Met Glu Thr Lys Met Lys Ile His Tyr Gln Val Ala Ala Val Val
 3965 3970 3975

ctg aca ggt gtt atg gtt tgg ggg ctt tcc cat tgg cgt tac acc gtc 14033
 Leu Thr Gly Val Met Val Trp Gly Leu Ser His Trp Arg Tyr Thr Val
 3980 3985 3990 3995

ggt tac cac gcg gca gat act caa tgg caa caa cgc cag gcc gaa cag 14081
 Gly Tyr His Ala Ala Asp Thr Gln Trp Gln Gln Arg Gln Ala Glu Gln
 4000 4005 4010

gaa agg gcc gat gcg ttg gcc ctc ctg gca gca gaa acc cgg gaa aga 14129

Glu Arg Ala Asp Ala Leu Ala Leu Leu Ala Ala Glu Thr Arg Glu Arg
 4015 4020 4025
 aag tgg gag cag caa cga cag act gac atg aac aag gtg gct ata cat 14177
 Lys Trp Glu Gln Gln Arg Gln Thr Asp Met Asn Lys Val Ala Ile His
 4030 4035 4040
 gct gaa gaa gaa ctg gct gct gcg cgt gac gct gcc gct gat gct cag 14225
 Ala Glu Glu Glu Leu Ala Ala Ala Arg Asp Ala Ala Ala Asp Ala Gln
 4045 4050 4055
 cgc act ggt cag cgc ctg cag cac acc gtt acc acc ctc cag cgg caa 14273
 Arg Thr Gly Gln Arg Leu Gln His Thr Val Thr Thr Leu Gln Arg Gln
 4060 4065 4070 4075
 ctt gcc agt cgt gaa acc cgc cgc ctt tcc gca gct acc gct atc ggt 14321
 Leu Ala Ser Arg Glu Thr Arg Arg Leu Ser Ala Ala Thr Ala Ile Gly
 4080 4085 4090
 aca gac gac ctc gga ggc caa ccc ggc gtt ttg ttt gcc gaa ctg ttc 14369
 Thr Asp Asp Leu Gly Gly Gln Pro Gly Val Leu Phe Ala Glu Leu Phe
 4095 4100 4105
 cgc cgc gct gac cag aga gcg gga gag ctg gca gcg tat gct gac agg 14417
 Arg Arg Ala Asp Gln Arg Ala Gly Glu Leu Ala Ala Tyr Ala Asp Arg
 4110 4115 4120
 acc aga gtg aaa tgg cag gcc tgc ggg cgc gcc tat cag gcg gct acg 14465
 Thr Arg Val Lys Trp Gln Ala Cys Gly Arg Ala Tyr Gln Ala Ala Thr
 4125 4130 4135
 cac gaa gca gaa aaa taa ggcgatttag ccgttaagga aaagtgcg 14513
 His Glu Ala Glu Lys
 4140 4145
 tgttttcgcg\ attaatatta acaggagatc ac atg agc aca tcc ttg ttc agt 14566
 Met Ser Thr Ser Leu Phe Ser
 4150
 agc acc ccg tcg gtc gcg gtg ctc gac aac cgc gcc ctg ttg gtg cgg 14614
 Ser Thr Pro Ser Val Ala Val Leu Asp Asn Arg Gly Leu Leu Val Arg
 4155 4160 4165
 gag ctg cag tac tac cgc cat ccg gat aca ccg gag gag acg gac gag 14662
 Glu Leu Gln Tyr Tyr Arg His Pro Asp Thr Pro Glu Glu Thr Asp Glu
 4170 4175 4180
 cgt atc acc tgc cat cag cac gat gag cgc gcc agc ttg tca caa agc 14710
 Arg Ile Thr Cys His Gln His Asp Glu Arg Gly Ser Leu Ser Gln Ser
 4185 4190 4195 4200
 gcc gac ccg cgg tta cac gcg gcc ggt ctg aca aat ttc acg tac ctg 14758
 Ala Asp Pro Arg Leu His Ala Ala Gly Leu Thr Asn Phe Thr Tyr Leu
 4205 4210 4215
 aat agc ctg acc ggg aca gta ctg cag agc gtc agc gcc gat gcc ggt 14806
 Asn Ser Leu Thr Gly Thr Val Leu Gln Ser Val Ser Ala Asp Ala Gly
 4220 4225 4230

acg tcg ctg gaa ctg agc gat gcc gcc ggg cgg gcg ttt ctg gcc gtc Thr Ser Leu Glu Leu Ser Asp Ala Ala Gly Arg Ala Phe Leu Ala Val 4235 4240 4245	14854
acc ggg gct ggg acg gaa gac gcg gtc acc cgc acc tgg caa tat gaa Thr Gly Ala Gly Thr Glu Asp Ala Val Thr Arg Thr Trp Gln Tyr Glu 4250 4255 4260	14902
gac gat acc ctg ccg ggc cgc ccg ctg agc atc acc gag cag gtt acc Asp Asp Thr Leu Pro Gly Arg Pro Leu Ser Ile Thr Glu Gln Val Thr 4265 4270 4275 4280	14950
ggc gaa gcc gcc caa att acg gaa cgc ttc gtg tac gct ggc aat acg Gly Glu Ala Ala Gln Ile Thr Glu Arg Phe Val Tyr Ala Gly Asn Thr 4285 4290 4295	14998
gat gcc gag aag att ctc aat ctg gct ggc cag tgt gtc agt cat tac Asp Ala Glu Lys Ile Leu Asn Leu Ala Gly Gln Cys Val Ser His Tyr 4300 4305 4310	15046
gat acc gcc gga ctg gtg cag acg gac agc atc gcc ctg agc ggc gtg Asp Thr Ala Gly Leu Val Gln Thr Asp Ser Ile Ala Leu Ser Gly Val 4315 4320 4325	15094
ccg ctc gcc gtc acg cgg cag ttg ctg ccc gac gcg gcg ggg gcc aac Pro Leu Ala Val Thr Arg Gln Leu Leu Pro Asp Ala Ala Gly Ala Asn 4330 4335 4340	15142
tgg atg ggt gag gat gcc tcg gcc tgg aat gac ctg ctg gat ggg gag Trp Met Gly Glu Asp Ala Ser Ala Trp Asn Asp Leu Leu Asp Gly Glu 4345 4350 4355 4360	15190
acg ttc ttc acc cag acc cac gct gat gcg acc ggc gcc gtc ctg agc Thr Phe Phe Thr Gln Thr His Ala Asp Ala Thr Gly Ala Val Leu Ser 4365 4370 4375	15238
atc acc gat gca aaa ggt aat ctg cag cgt gtg gca tat gat gtg gct Ile Thr Asp Ala Lys Gly Asn Leu Gln Arg Val Ala Tyr Asp Val Ala 4380 4385 4390	15286
ggg ctg cta tcg ggc agt tgg ttg acg ctg aag gac ggc acg gag cag Gly Leu Leu Ser Gly Ser Trp Leu Thr Leu Lys Asp Gly Thr Glu Gln 4395 4400 4405	15334
gtc atc gtg gcc tcc ctg acg tac tcg gcc gcc ggg aaa aag ttg cgt Val Ile Val Ala Ser Leu Thr Tyr Ser Ala Ala Gly Lys Lys Leu Arg 4410 4415 4420	15382
gaa gaa cac ggc aac ggc gtg gta acc tcg tat att tac gag ccg gaa Glu Glu His Gly Asn Gly Val Val Thr Ser Tyr Ile Tyr Glu Pro Glu 4425 4430 4435 4440	15430
aca cag cgc ctg acg ggg att aaa acg gaa cgt ccg tct ggg cac gtt Thr Gln Arg Leu Thr Gly Ile Lys Thr Glu Arg Pro Ser Gly His Val 4445 4450 4455	15478
gcc gga gca aaa gtg ctg cag gac ctg cgc tat acg tat gac ccg gta	15526

Ala Gly Ala Lys Val Leu Gln Asp Leu Arg Tyr Thr Tyr Asp Pro Val	
4460 4465 4470	
ggc aac gta ctc agc gtc aat aac gat gcg gaa gag acc cgc ttc tgg	15574
Gly Asn Val Leu Ser Val Asn Asn Asp Ala Glu Glu Thr Arg Phe Trp	
4475 4480 4485	
cgt aac cag aaa gtg gta ccg gag aat acg tac atc tac gac agc ctg	15622
Arg Asn Gln Lys Val Val Pro Glu Asn Thr Tyr Ile Tyr Asp Ser Leu	
4490 4495 4500	
tac cag ctg gtc agc gcc aca ggg cgt gag atg gcc aat gcc ggc cag	15670
Tyr Gln Leu Val Ser Ala Thr Gly Arg Glu Met Ala Asn Ala Gly Gln	
4505 4510 4515 4520	
cag ggc aac gac tta cca tcc gct aca gcc ccc ctt cct aca gac agc	15718
Gln Gly Asn Asp Leu Pro Ser Ala Thr Ala Pro Leu Pro Thr Asp Ser	
4525 4530 4535	
tct gcc tac acc aat tac acg cgc acc tac cgt tat gac cgt ggt ggc	15766
Ser Ala Tyr Thr Asn Tyr Thr Arg Thr Tyr Arg Tyr Asp Arg Gly Gly	
4540 4545 4550	
aac ctg acg cag atg cgc cac agt gcc cct gcc acg aac aat aat tat	15814
Asn Leu Thr Gln Met Arg His Ser Ala Pro Ala Thr Asn Asn Asn Tyr	
4555 4560 4565	
acg aca gac atc acg gtt agt gac cgc agc aat agg gcg gta ctg agc	15862
Thr Thr Asp Ile Thr Val Ser Asp Arg Ser Asn Arg Ala Val Leu Ser	
4570 4575 4580	
acg ttg gcg gaa gtg ccg tca gat gtt gat atg ctg ttc agt gca gga	15910
Thr Leu Ala Glu Val Pro Ser Asp Val Asp Met Leu Phe Ser Ala Gly	
4585 4590 4595 4600	
ggt cac cag aag cac ctg cag ccg ggg caa gca ctg gtg tgg acg cca	15958
Gly His Gln Lys His Leu Gln Pro Gly Gln Ala Leu Val Trp Thr Pro	
4605 4610 4615	
cgt gga gaa ctg caa aag gtg aca ccg gtg gtg cgt gat ggg ggg gcg	16006
Arg Gly Glu Leu Gln Lys Val Thr Pro Val Val Arg Asp Gly Gly Ala	
4620 4625 4630	
gac gac agc gaa agc tat cgg tat gat gcg ggc agt cag cgt att atc	16054
Asp Asp Ser Glu Ser Tyr Arg Tyr Asp Ala Gly Ser Gln Arg Ile Ile	
4635 4640 4645	
aaa acc ggc acg cgg caa act ggc aac aac gtt cag aca cag cgg gta	16102
Lys Thr Gly Thr Arg Gln Thr Gly Asn Asn Val Gln Thr Gln Arg Val	
4650 4655 4660	
gtg tac ctg ccg ggg ctg gag tta cgt atc atg gca aat ggc gtg acg	16150
Val Tyr Leu Pro Gly Leu Glu Leu Arg Ile Met Ala Asn Gly Val Thr	
4665 4670 4675 4680	
gaa aaa gaa agc ctg cag gtt att acg gtg ggc gag gct ggg cgg gca	16198
Glu Lys Glu Ser Leu Gln Val Ile Thr Val Gly Glu Ala Gly Arg Ala	
4685 4690 4695	

caa gtg cgc gta ttg cac tgg gag atc ggc aag ccg gat gac ctc gat Gln Val Arg Val Leu His Trp Glu Ile Gly Lys Pro Asp Asp Leu Asp 4700 4705 4710	16246
gag gac tgc gtg cgt tac agt tac gat aac ctg gtg ggc agc agc cag Glu Asp Ser Val Arg Tyr Ser Tyr Asp Asn Leu Val Gly Ser Ser Gln 4715 4720 4725	16294
ctg gag ctg gac aga gag ggt tac ctt atc agt gag gag gag ttc tac Leu Glu Leu Asp Arg Glu Gly Tyr Leu Ile Ser Glu Glu Glu Phe Tyr 4730 4735 4740	16342
ccg tat ggc gga acg gct gtt ctg acg gcg cga agt gag gtt gag gct Pro Tyr Gly Gly Thr Ala Val Leu Thr Ala Arg Ser Glu Val Glu Ala 4745 4750 4755 4760	16390
gac tac aaa act atc cga tac tca ggc aag gag cgt gac gcg acg ggg Asp Tyr Lys Thr Ile Arg Tyr Ser Gly Lys Glu Arg Asp Ala Thr Gly 4765 4770 4775	16438
ctg gat tat tac ggt tat cgg tat tac cag cca tgg gca ggg cgc tgg Leu Asp Tyr Tyr Gly Tyr Arg Tyr Tyr Gln Pro Trp Ala Gly Arg Trp 4780 4785 4790	16486
ctc tcc acg gac ccg gca ggc acg gtg gac ggg ctg aac ctg ttc cgc Leu Ser Thr Asp Pro Ala Gly Thr Val Asp Gly Leu Asn Leu Phe Arg 4795 4800 4805	16534
atg gtg cgg aat aat ccc gtc acg ctg ttt gac agc aac ggg cgg atc Met Val Arg Asn Asn Pro Val Thr Leu Phe Asp Ser Asn Gly Arg Ile 4810 4815 4820	16582
agt act ggt cag gag gcc aga cga tta gtg ggg gaa gca ttt gtt cat Ser Thr Gly Gln Glu Ala Arg Arg Leu Val Gly Glu Ala Phe Val His 4825 4830 4835 4840	16630
ccg tta cac atg cct gtt ttt gaa aga att tct gta gag aga aag att Pro Leu His Met Pro Val Phe Glu Arg Ile Ser Val Glu Arg Lys Ile 4845 4850 4855	16678
tca atg agc gta agg gaa gct ggc att tat act att tca gcg ctg ggt Ser Met Ser Val Arg Glu Ala Gly Ile Tyr Thr Ile Ser Ala Leu Gly 4860 4865 4870	16726
gaa ggt gca gca gca aaa ggc cat aat att cta gag aaa acc att aaa Glu Gly Ala Ala Ala Lys Gly His Asn Ile Leu Glu Lys Thr Ile Lys 4875 4880 4885	16774
ccc ggt tcc ctg aag gct atc tat ggt gat aaa gct gag tca att ctt Pro Gly Ser Leu Lys Ala Ile Tyr Gly Asp Lys Ala Glu Ser Ile Leu 4890 4895 4900	16822
gga ctg gca aaa cgt agc ggt ctc gtt ggc cga gta gga cag tgg gat Gly Leu Ala Lys Arg Ser Gly Leu Val Gly Arg Val Gly Gln Trp Asp 4905 4910 4915 4920	16870
gca tca ggt gta cgt gga att tat gcg cac aac aga ccg ggt ggt gag	16918

Ala Ser Gly Val Arg Gly Ile Tyr Ala His Asn Arg Pro Gly Gly Glu
 4925 4930 4935

gat ttg gtt tat cct gtc agc ctg cag aat act tct gcc aat gaa att 16966
 Asp Leu Val Tyr Pro Val Ser Leu Gln Asn Thr Ser Ala Asn Glu Ile
 4940 4945 4950

gtt aat gca tgg ata aaa ttt aaa atc atc acg ccc tac acc ggg gat 17014
 Val Asn Ala Trp Ile Lys Phe Lys Ile Ile Thr Pro Tyr Thr Gly Asp
 4955 4960 4965

tat gac atg cac gat att att aaa ttc tct gat ggg aaa ggg cat gtg 17062
 Tyr Asp Met His Asp Ile Ile Lys Phe Ser Asp Gly Lys Gly His Val
 4970 4975 4980

cct aca gcg gaa agt agt gag gaa aga gga gta aaa gat cta att aat 17110
 Pro Thr Ala Glu Ser Ser Glu Glu Arg Gly Val Lys Asp Leu Ile Asn
 4985 4990 4995 5000

aaa ggt gtt gcg gag gtc gat cct tcc aga ccc ttt gag tat aca gcg 17158
 Lys Gly Val Ala Glu Val Asp Pro Ser Arg Pro Phe Glu Tyr Thr Ala
 5005 5010 5015

atg aat gtt att cgc cat gga cca cag gtg aac ttt gtt ccc tat atg 17206
 Met Asn Val Ile Arg His Gly Pro Gln Val Asn Phe Val Pro Tyr Met
 5020 5025 5030

tgg gaa cat gag cac gat aaa gtc gtt aat gat aat ggt tat ctg ggg 17254
 Trp Glu His Glu His Asp Lys Val Val Asn Asp Asn Gly Tyr Leu Gly
 5035 5040 5045

gtg gta gct agc ccg ggg ccg ttc ccg gta gcg atg gta cat cag ggg 17302
 Val Val Ala Ser Pro Gly Pro Phe Pro Val Ala Met Val His Gln Gly
 5050 5055 5060

gaa tgg act gtt ttt gac aac agt gaa gaa ctg ttt aat ttc tat aaa 17350
 Glu Trp Thr Val Phe Asp Asn Ser Glu Leu Phe Asn Phe Tyr Lys
 5065 5070 5075 5080

tct aca aat aca cct ctt cct gaa cac tgg tcc caa gat ttt atg gac 17398
 Ser Thr Asn Thr Pro Leu Pro Glu His Trp Ser Gln Asp Phe Met Asp
 5085 5090 5095

aga ggg aaa gga ata gtc gca act cct cgg cat gct gaa ctt ctt gat 17446
 Arg Gly Lys Gly Ile Val Ala Thr Pro Arg His Ala Glu Leu Leu Asp
 5100 5105 5110

aaa cga cga gtc atg tac taa tcgtaacgat ttctgcctt acccaaagta 17497
 Lys Arg Arg Val Met Tyr
 5115

tacagcccg tgagacattt tctctgtctc atttgggttg tttttgtctc atctgcatgt 17557

tatgtcttcc ctcattctaaa gtctaacgag acatttttag caaaatggca ctttacgggt 17617

atgttcgcgt ttcaaccgac ggtccggatt ttactctgta aatacagaca ctctgcgcag 17677

cctgctgcga aattatccgt gcgaaaaaag ccagcggcag cagccgggat ggacgaaatg 17737

aactgcagct tctgctggct tttttgcggc caggcaacat gctgatgggt acgtgagttg 17797
atcggctgcc accaaaaagt ccggagcgtg cggcccagat cgccgcaata ataactgctgt 17857
atggatattc catcaccact gtatatcgca cactctgggc cttccagaaa ccccataccg 17917
cacaccggtg tgatcgctgg aagccccggg cattaccgcc gtctgtactc gaacactatt 17977
gtggacttga tggtaggag attgaatcga ccatttttga gatccctaac catagatcgt 18037
agagttgcac actcccagat ggctgggtt agcagcgtat tatgcttaa aattcatgtt 18097
ttgctgtgtt tttaatccaa aacctgcttt tcaggcgac ttatccagct acgggggtctg 18157
aagccatcgt tttttgccc tacgatgtag cctgtcagag agcatttttg tggcgtgctc 18217
gcccgtacg gtaccggcgg caaacgcag ccggcctttg cagaggatgc actggtacgg 18277
atcgggtccc aggaagcctt tcacagcac cgcgaaccgg ggccgtttcg gtttctccc 18337
taccgtcatc tccagcgt cgtaaaccct cggcagcagc gtgcccgttt gcggttgcc 18397
agaaaacat agtaacgcac cattttaaaa tgcgtgcag ggatatggct gacgtaacgc 18457
tgcagcatct cctcctggct gattttctgg cgtttgtgct gctgcgtacg gtgatcgtaa 18517
tactgatgca ccacggcccc gccgcggtag tggcgtagct gagaagccgc caccggcggg 18577
cgcttcaggt accgggccag gtatttcacg ctgcgccagg cgccgcgggt ctttttgca 18637
aaattcactt tccagggcg gcggtattgc gcatgcaggg tcttcgttgc ggatatggcc 18697
gagaccggc agggcgccag gattgatgc cagcaggta acgacggcat tgcgccagat 18757
ggcttccacc tctttctttt taaagaacag ctgccgccag acgtgggtgt tgacgtcaag 18817
accgccggg gtaacggaga cgtggatatg cggatgttga ttgagctgcc ggccttaggt 18877
gtggagcgc caaaaaatgc cggcctcgat gcctgccgg cgtgcccgagg ggagcatggc 18937

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 144 amino acid residues
- (B) TYPE: amino acid
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: PROTEIN (ORF 1)

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

Met Lys Ile Ser Ser Arg Gly Ile Ala Leu Ile Lys Glu Phe Glu Gly
 1             5             10             15

Leu Arg Leu His Ala Tyr Arg Cys Ala Ala Asp Val Trp Thr Val Gly
      20             25             30

Tyr Gly His Thr Ala Gly Val Thr Lys Gly Asp Ile Ile Thr Val Asp
      35             40             45

Glu Ala Gln Thr Met Leu Thr Asn Asp Ile Thr Val Phe Glu Arg Ala
      50             55             60

Val Ser Gln Ala Val Ala Val Pro Leu Asn Gln Ser Gln Tyr Asp Ala
      65             70             75             80

Leu Val Ser Leu Val Phe Asn Ile Gly Gln Gly Asn Phe Lys Arg Ser
      85             90             95

Thr Leu Leu Lys Lys Leu Asn Lys Gln Asp Tyr Val Gly Ala Gly Asn
      100            105            110

Glu Phe Leu Arg Trp Thr Arg Ala Asn Gly Lys Val Leu Pro Gly Leu
      115            120            125

Ile Arg Arg Arg Glu Ala Glu Arg Val Leu Phe Glu Lys Leu Gly
      130            135            140

Ala

```

(2) INFORMATION FOR SEQ ID NO: 3:

- (i). SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 191 amino acid residues
 (B) TYPE: amino acid
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: PROTEIN (ORF 2)

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

Met Ser Pro Ser Pro Leu Thr Gly Ala Ala Leu Met Glu Thr Lys Met
  1             5             10             15
Lys Ile His Tyr Gln Val Ala Ala Val Val Leu Thr Gly Val Met Val
      20             25             30
Trp Gly Leu Ser His Trp Arg Tyr Thr Val Gly Tyr His Ala Ala Asp
      35             40             45
Thr Gln Trp Gln Gln Arg Gln Ala Glu Gln Glu Arg Ala Asp Ala Leu
      50             55             60
Ala Leu Leu Ala Ala Glu Thr Arg Glu Arg Lys Trp Glu Gln Gln Arg
      65             70             75             80
Gln Thr Asp Met Asn Lys Val Ala Ile His Ala Glu Glu Glu Leu Ala
      85             90             95
Ala Ala Arg Asp Ala Ala Ala Asp Ala Gln Arg Thr Gly Gln Arg Leu
      100            105            110
Gln His Thr Val Thr Thr Leu Gln Arg Gln Leu Ala Ser Arg Glu Thr
      115            120            125
Arg Arg Leu Ser Ala Ala Thr Ala Ile Gly Thr Asp Asp Leu Gly Gly
      130            135            140
Gln Pro Gly Val Leu Phe Ala Glu Leu Phe Arg Arg Ala Asp Gln Arg
      145            150            155            160
Ala Gly Glu Leu Ala Ala Tyr Ala Asp Arg Thr Arg Val Lys Trp Gln
      165            170            175
Ala Cys Gly Arg Ala Tyr Gln Ala Ala Thr His Glu Ala Glu Lys
      180            185            190

```

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2376 amino acid residues
 (B) TYPE: amino acid
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: PROTEIN (SepA)

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

Met Arg Gln Asp Ile Met Tyr Asn Ile Asp Asp Ile Leu Glu Lys Val
 1             5             10             15
Asn Ala Pro Arg Ala Arg Leu Ser Glu Glu Asn Asp Thr Ala Val Thr
      20             25             30
Leu Thr Asp Leu Phe Ser Arg Ser Phe Pro Glu Val Lys Lys Ile Thr
      35             40             45
Gly Asp Ser Leu Ser Trp Gly Glu Val Cys Tyr Leu Tyr Ser Gln Ala
      50             55             60
Gln His Glu Gln Lys Glu Asn Arg Leu Thr Glu Ser Arg Ile Leu Ala
      65             70             75             80
Arg Ala Asn Pro Leu Leu Val Asn Ala Val Arg Leu Gly Ile Arg Gln
      85             90             95
Ala Ala Gly Ser Arg Ser Tyr Asp Asp Trp Phe Gly Ser Arg Ala Asp
      100            105            110
Arg Phe Ala Arg Pro Gly Ser Val Ala Ser Met Phe Ser Pro Ala Ala
      115            120            125
Tyr Leu Thr Glu Leu Tyr Arg Glu Ala Lys Asp Leu His Pro Asp Thr
      130            135            140
Ser Leu Phe Arg Leu Asp Ile Arg Arg Pro Asp Leu Ala Ala Leu Ala
      145            150            155            160
Leu Ser Gln Asn Asn Met Asp Asp Glu Leu Ser Thr Leu Ser Leu Ser
      165            170            175
Asn Glu Leu Leu Tyr Arg Gly Ile Gly Ala Ala Glu Gly Leu Asp Asp
      180            185            190

```

Asp Ser Val Arg Glu Leu Leu Ala Gly Tyr Arg Leu Thr Gly Leu Thr
 195 200 205
 Pro Tyr His Trp Ala Tyr Glu Ala Ala Arg Gln Ala Ile Leu Val Gln
 210 215 220
 Asp Pro Thr Leu Met Gly Phe Ser Arg Asn Pro Asp Val Ala Gln Leu
 225 230 235 240
 Met Asp Pro Ala Ser Met Leu Ala Ile Glu Ala Asp Ile Ser Pro Glu
 245 250 255
 Leu Tyr Gln Ile Leu Ala Glu Glu Ile Thr Thr Asp Ser Tyr Glu Ala
 260 265 270
 Leu Trp Ser Lys Asn Phe Gly Asp Met Pro Pro Ser Ser Leu Leu Ser
 275 280 285
 Tyr Asp Ala Leu Ala Thr Phe Tyr Asp Leu Asp Tyr Asp Glu Leu Thr
 290 295 300
 Ser Leu Leu Ser Leu Arg Leu Asp Phe Ser Asn Pro Asn Asn Glu Tyr
 305 310 315 320
 Tyr Ile Asn Ser Gln Leu Ser Val Val Thr Leu Asn Glu Ser Thr Gly
 325 330 335
 Leu Ile Thr Ile His His Tyr Leu Arg Thr Leu Gly Gly Asp Ser Gln
 340 345 350
 Gln Ile Asn Pro Glu Leu Ile Pro Tyr Gly Asp Gly Thr Tyr Leu Tyr
 355 360 365
 Asn Phe Ser Val Val Ser Thr Ile Ser Glu Asp Ser Phe Lys Leu Gly
 370 375 380
 Ser Leu Gly Ser Asn Ser Ser Asn Leu Tyr Ser Gly Asp Tyr Gln Leu
 385 390 395 400
 Gln Lys Gly Val Arg Tyr Ser Ile Pro Val Glu Ile Asp Glu Gly Lys
 405 410 415
 Leu Asn Asp Gly Ile Thr Ile Gly Leu Ser Arg Lys Gly Gly Tyr
 420 425 430
 Tyr Ser Thr Val Asn Phe Thr Leu Ile Glu Tyr Asp Pro Ala Ile Phe
 435 440 445
 Ile Leu Lys Leu Asn Lys Val Ile Arg Leu Tyr Lys Ala Thr Gly Met
 450 455 460
 Thr Thr Ala Glu Ile Tyr Gln Ile Thr Asn Ile Leu Asn Asn Gly Leu
 465 470 475 480
 Thr Ile Asp His Ala Val Leu Ser Lys Ile Phe Leu Val Arg Tyr Leu
 485 490 495
 Met Arg His Tyr Gln Leu Asp Val Ala Arg Ser Leu Ile Leu Cys Asn

[illegible]

Ser His Ala Gly Glu Val Leu Thr Ala Leu Glu Thr Gly Glu Leu Ser
 820 825 830
 Ser Ala Leu Leu Ala Arg Ala Leu Ser Gln Asn Glu Gln Asp Val Thr
 835 840 845
 Gly Ala Leu Ala Gln Val Arg Gly Ala Gly Glu Gln Asp Asn Ser Val
 850 855 860
 Phe Thr Ser Trp Glu Glu Val Asp Gln Ala Glu Gln Trp Leu Asp Met
 865 870 875 880
 Ser Glu Thr Leu Ser Ile Thr Pro Ser Gly Leu Ala Ser Leu Ile Ala
 885 890 895
 Leu Lys Tyr Ile Asn Val Ser Asp Asp Ser Ala Pro Leu Tyr Ser Gln
 900 905 910
 Trp Gln Val Val Ser Gly Leu Leu Gln Ala Gly Leu Lys Ser Ser Gln
 915 920 925
 Ser Ser Ala Leu His Asp Tyr Leu Glu Glu Gly Thr Ser Ser Ala Leu
 930 935 940
 Cys Ala Tyr Tyr Leu Arg Asn Leu Ala Pro Asn Met Val Ser Gly Arg
 945 950 955 960
 Asp Asp Leu Phe Gly Tyr Leu Leu Leu Asp Asn Gln Val Ser Ala Lys
 965 970 975
 Val Lys Thr Thr Arg Ile Ala Glu Ala Ile Ala Gly Ile Arg Leu Tyr
 980 985 990
 Ile Asn Arg Ala Leu Asn Gly Ile Glu Leu Ser Ala Met Ala Glu Val
 995 1000 1005
 Arg Gly Arg Gln Phe Phe Thr Asp Trp Asp Thr Phe Asn Lys Arg Tyr
 1010 1015 1020
 Ser Thr Trp Ala Gly Val Ser Glu Leu Val Tyr Tyr Pro Glu Asn Tyr
 1025 1030 1035 1040
 Leu Asp Pro Thr Val Arg Ile Gly Gln Thr Gly Met Met Asp Thr Leu
 1045 1050 1055
 Leu Gln Ser Val Ser Gln Ser Ser Ile Asn Arg Asp Thr Val Glu Asp
 1060 1065 1070
 Ala Phe Lys Thr Tyr Leu Thr Thr Phe Glu Gln Ile Ala Asn Leu Asn
 1075 1080 1085
 Thr Val Ser Gly Tyr His Asp Asn Ala Ser Met Thr Gln Gly Thr Thr
 1090 1095 1100
 Trp Tyr Val Gly Arg Ser Ile Thr Asp Gln Thr Asn Trp Tyr Trp Arg
 1105 1110 1115 1120

Ser Ala Asn His Ser Lys Ile Gln Asp Ser Met Met Pro Ala Asn Ala
 1125 1130 1135
 Trp Thr Gly Trp Thr Lys Ile Asn Cys Gly Met Asn Pro Trp Ser Asp
 1140 1145 1150
 Leu Val Cys Ser Val Phe Phe Asn Ser Arg Leu Tyr Val Val Trp Val
 1155 1160 1165
 Glu Glu Asn Gln Ser Ala Asp Thr Glu Ala Glu Ser Thr Thr Thr Thr
 1170 1175 1180
 Gln Gln Ser Tyr Thr Leu Lys Leu Ser Phe Arg Arg Tyr Asp Gly Thr
 1185 1190 1195 1200
 Trp Ser Ser Pro Val Ser Phe Asp Ile Thr Gly Asn Ile Ala Phe Pro
 1205 1210 1215
 Glu Thr Gln Gly Met His Val Thr Cys Asn Pro Leu Thr Glu Gln Leu
 1220 1225 1230
 Tyr Cys Ala Phe Tyr Ser Val Thr Ser Lys Pro Asp Phe Asp Asn Ala
 1235 1240 1245
 Gln Leu Ile Ser Val Asp Asn Asp Met Thr Leu Asn Val Ile Ser Asp
 1250 1255 1260
 Ile Gly Ile Phe Lys Ser Val Ser His Glu Phe Asn Thr Ser Thr Glu
 1265 1270 1275 1280
 Lys Phe Ile Asn Asn Val Phe Ser Asp Pro Ser Ala Asn Tyr Phe Val
 1285 1290 1295
 Ser Ala Thr Ser Ser Leu Ile Asp Asp Val Ile His Ser Asp Phe Ser Leu
 1300 1305 1310
 Leu Asn Ser Lys Thr Thr Ser Thr Val Phe Thr Asn Glu Asp Ser Ser
 1315 1320 1325
 Leu Leu Thr Pro Glu Leu His Ile Thr Ala Asn Val Ser Cys Phe Val
 1330 1335 1340
 Ser Thr Ala Gly Ile Ala Thr Gln Ser Thr Ile Glu Lys Phe Val Gln
 1345 1350 1355 1360
 Ala Gly Ile Glu Phe Glu Glu Ile Asn Phe Tyr Ala Gly Gln Ala Ala
 1365 1370 1375
 Gly Gly Phe Asp Gly Phe Val Gly Val Asp Val Ser Asn Ser Lys Val
 1380 1385 1390
 Tyr Gln Val Gly Lys Glu Ala Val Gly Val Thr Val Lys Ser Tyr Ser
 1395 1400 1405
 Val Thr Gly Val Ser Gly Ser Val Glu Leu Phe Ile Asp Ser Ser Asn
 1410 1415 1420
 Lys Tyr Phe Ser Gly Ile Leu Ser Asp Lys Met Ile Thr Ala Leu Ile

425 1430 1435 1440
 Ser Gly Ser Thr Ser Lys Val Asn Tyr Val Ser Ser Ile Gly Ser Gln
 1445 1450 1455
 Asp Phe Trp Ser Val Lys Ser Leu Met Pro Ala Leu Gln Ile Tyr Glu
 1460 1465 1470
 Leu Ile Asp Asp Ile Ile Leu Thr Ser Gly Val Asn Gly Thr Glu Ile
 1475 1480 1485
 Lys Ser Trp Pro Ser Ala Glu Trp Tyr Asn Asp Lys Leu Ser Leu Gln
 1490 1495 1500
 Ser Gly Asn Asn Leu Phe Asn Thr Lys Ser Leu Ser Phe Thr Val Asn
 505 1510 1515 1520
 Thr Ser Asp Ile Val Glu Asp Glu Phe Asp Val Thr Phe Thr Phe Thr
 1525 1530 1535
 Ala Val Asp Gln Asn Asn Val Val Leu Ala Ala Arg Thr Ala Ile Leu
 1540 1545 1550
 Thr Val Ile Arg Asn Ile Asn Asn Asp Thr Ser Val Ile Ala Leu Arg
 1555 1560 1565
 Lys Asn Thr Arg Gly Ala Gln Tyr Ile Arg Phe Thr Ala Gly Asn Asp
 1570 1575 1580
 Val Ala Leu Ile Arg Leu Asn Thr Leu Phe Ala Arg Gln Leu Val Asp
 585 1590 1595 1600
 Arg Ala Asn Thr Gly Ile Asp Thr Ile Leu Ser Met Glu Thr Gln Arg
 1605 1610 1615
 Leu Thr Glu Pro Ala Leu Glu Glu Gly Ser Asp Val Phe Met Asp Phe
 1620 1625 1630
 Ser Gly Ala Asn Ala Leu Tyr Phe Trp Glu Leu Phe Tyr Tyr Thr Pro
 1635 1640 1645
 Met Met Val Phe Gln Arg Leu Leu Gln Glu Gln His Phe Pro Glu Ala
 1650 1655 1660
 Thr Arg Trp Leu Gln Tyr Val Trp Asn Pro Ala Gly His Val Val Asn
 665 1670 1675 1680
 Gly Val Leu Gln Asn Tyr Thr Trp Asn Val Arg Pro Leu Glu Glu Asp
 1685 1690 1695
 Thr Gly Trp Asn Asp Ser Pro Leu Asp Ser Ile Asp Pro Asp Ala Ile
 1700 1705 1710
 Ala Gln Tyr Asp Pro Met His Tyr Lys Val Ala Thr Phe Met Ser Tyr
 1715 1720 1725
 Leu Asp Leu Leu Ile Ala Arg Gly Asp Ala Ala Tyr Arg Leu Leu Glu
 1730 1735 1740

Arg Asp Thr Leu Asn Glu Ala Arg Met Trp Tyr Val Gln Ala Leu Asn
 745 1750 1755 1760
 Leu Leu Gly Asp Glu Pro Tyr Ile Ser Phe Asp Ala Asp Trp Ser Ala
 1765 1770 1775
 Leu Thr Leu Gly Asp Ala Ala Ser Glu Val Thr Arg Arg Asp Tyr Gln
 1780 1785 1790
 Glu Ala Leu Leu Ala Val Arg Arg Leu Val Pro Ala Pro Glu Thr Arg
 1795 1800 1805
 Thr Ala Asn Ser Leu Thr Ala Leu Phe Leu Pro Gln Gln Asn Glu Val
 1810 1815 1820
 Leu Lys Gly Tyr Trp Gln Thr Leu Ala Gln Arg Leu His Asn Leu Arg
 825 1830 1835 1840
 His Asn Leu Ser Ile Asp Gly Gln Pro Leu Ser Leu Ser Val Tyr Ala
 1845 1850 1855
 Thr Pro Ser Glu Pro Ser Ala Leu Gln Ser Ala Val Val Asn Ser Ala
 1860 1865 1870
 Gln Gly Ala Ala Ala Leu Pro Ala Ala Val Met Pro Leu Tyr Ser Phe
 1875 1880 1885
 Pro Val Met Leu Glu Asn Ala Arg Gly Met Val Ser Leu Leu Thr Gly
 1890 1895 1900
 Phe Gly Asn Thr Leu Leu Gly Ile Thr Glu Arg Gln Asp Ala Glu Ala
 905 1910 1915 1920
 Leu Ala Lys Leu Leu Gln Thr Gln Gly Ser Glu Leu Ile Arg Gln Gly
 1925 1930 1935
 Leu Arg Gln Gln Asp Asn Val Leu Glu Glu Ile Asp Ala Asp Ile Ala
 1940 1945 1950
 Ala Leu Glu Glu Ser Arg Arg Gly Ala Gln Met Arg Phe Glu Arg Tyr
 1955 1960 1965
 Lys Val Leu Tyr Glu Ala Asp Val Asn Thr Gly Glu Lys Gln Ala Met
 1970 1975 1980
 Asp Leu Tyr Leu Ser Ser Ser Val Leu Ser Ala Ser Thr Ala Ala Leu
 985 1990 1995 2000
 Phe Leu Ala Glu Ala Ala Ala Asp Met Leu Pro Asn Ile Tyr Gly Leu
 2005 2010 2015
 Ala Val Gly Gly Ser Arg Tyr Gly Ala Leu Phe Lys Ala Thr Ala Ile
 2020 2025 2030
 Gly Ile Gln Val Ser Ser Asp Ala Thr Arg Ile Ser Ala Asp Lys Ile
 2035 2040 2045

Ser Gln Ser Glu Val Tyr Arg Arg Arg Arg Glu Glu Trp Glu Ile Gln
 2050 2055 2060
 Arg Asp Ser Ala Gln Ser Asp Val Ala Gln Ile Asp Ala Gln Leu Ala
 065 2070 2075 2080
 Ala Met Ala Val Arg Arg Glu Gly Ala Glu Leu Gln Lys Thr Tyr Leu
 2085 2090 2095
 Glu Thr Gln Gln Thr Gln Ala Gln Ala Gln Leu Ala Phe Leu Gln Ser
 2100 2105 2110
 Lys Phe Asn Asn Thr Ala Leu Tyr Ser Trp Leu Arg Gly Arg Leu Ser
 2115 2120 2125
 Ala Ile Tyr Tyr Gln Phe Tyr Asp Leu Ala Val Ser Arg Cys Leu Met
 2130 2135 2140
 Ala Gln Gln Ala Trp Gln Trp Asp Lys Phe Glu Thr Arg Ser Phe Ile
 145 2150 2155 2160
 Gln Pro Gly Ala Trp Met Gly Ala Asn Ala Gly Leu Leu Ala Gly Glu
 2165 2170 2175
 Thr Leu Met Leu Asn Leu Ala Gln Met Glu Gln Ala Trp Leu Thr Gly
 2180 2185 2190
 Asp Glu Arg Ala Ile Glu Val Thr Arg Thr Val Cys Leu Ser Glu Val
 2195 2200 2205
 Tyr Thr Ser Leu Ala Glu Asp Ala Ala Phe Ser Leu Ala Asp Lys Val
 2210 2215 2220
 Val Glu Leu Val Ser Asn Gly Ser Gly Ser Ala Gly Thr Lys Ser Asn
 225 2230 2235 2240
 Gly Leu Gln Met Asp Gln Gln Gln Leu Glu Ala Thr Leu Lys Leu Ala
 2245 2250 2255
 Asp Leu Gly Ile Gly Asn Asp Tyr Pro Val Ser Leu Gly Thr Met Arg
 2260 2265 2270
 Arg Ile Lys Gln Ile Ser Val Thr Leu Pro Ala Leu Val Gly Pro Tyr
 2275 2280 2285
 Gln Asp Val Arg Ala Val Leu Ser Tyr Gly Gly Ser Met Val Met Pro
 2290 2295 2300
 Arg Gly Cys Ser Ala Leu Ala Val Ser His Gly Met Asn Asp Ser Gly
 305 2310 2315 2320
 Gln Phe Gln Leu Asp Phe Asn Asp Pro Arg Tyr Leu Pro Phe Glu Gly
 2325 2330 2335
 Leu Pro Val Asp Asp Thr Gly Thr Leu Thr Leu Ser Phe Pro Asp Ala
 2340 2345 2350
 Asp Gly Lys Gln Gln Ala Met Leu Leu Ser Leu Ser Asp Ile Ile Leu
 2355 2360 2365
 His Ile Arg Tyr Thr Ile Ile Ser
 2370 2375

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1429 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: PROTEIN (SepB)

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Met Gln Asn His Gln Asp Met Ala Ile Thr Ala Pro Thr Leu Pro Ser
1 5 10 15
Gly Gly Gly Ala Val Thr Gly Leu Lys Gly Asp Ile Ala Ala Ala Gly
20 25 30
Pro Asp Gly Ala Ala Thr Leu Ser Ile Pro Leu Pro Val Ser Pro Gly
35 40 45
Arg Gly Tyr Ala Pro Thr Gly Ala Leu Asn Tyr His Ser Arg Ser Gly
50 55 60

Asn Gly Pro Phe Gly Ile Gly Trp Gly Ile Gly Gly Ala Ala Val Gln
 65 70 75 80
 Arg Arg Thr Arg Asn Gly Ala Pro Thr Tyr Asp Asp Thr Asp Glu Phe
 85 90 95
 Thr Gly Pro Asp Gly Glu Val Leu Val Pro Ala Leu Thr Ala Ala Gly
 100 105 110
 Thr Gln Glu Ala Arg Gln Ala Thr Ser Leu Leu Gly Ile Asn Pro Gly
 115 120 125
 Gly Ser Phe Asn Val Gln Val Tyr Arg Ser Arg Thr Glu Gly Ser Leu
 130 135 140
 Ser Arg Leu Glu Arg Trp Leu Pro Ala Asp Glu Thr Glu Thr Glu Phe
 145 150 155 160
 Trp Val Leu Tyr Thr Pro Asp Gly Gln Val Ala Leu Leu Gly Arg Asn
 165 170 175
 Ala Gln Ala Arg Ile Ser Asn Pro Thr Ala Pro Thr Gln Thr Ala Val
 180 185 190
 Trp Leu Met Glu Ser Ser Val Ser Leu Thr Gly Glu Gln Met Tyr Tyr
 195 200 205
 Gln Tyr Arg Ala Glu Asp Asp Asp Gly Cys Asp Glu Ala Glu Arg Asp
 210 215 220
 Ala His Pro Gln Ala Gly Ala Gln Arg Tyr Pro Val Ala Val Trp Tyr
 225 230 235 240
 Gly Asn Arg Gln Ala Ala Arg Thr Leu Pro Ala Leu Val Ser Thr Pro
 245 250 255
 Ser Met Asp Ser Trp Leu Phe Ile Leu Val Phe Asp Tyr Gly Glu Arg
 260 265 270
 Ser Ser Val Leu Ser Glu Ala Pro Ala Trp Gln Thr Pro Gly Ser Gly
 275 280 285
 Glu Trp Leu Cys Arg Gln Asp Cys Phe Ser Gly Tyr Glu Phe Gly Phe
 290 295 300
 Asn Leu Arg Thr Arg Arg Leu Cys Arg Gln Val Leu Met Phe His Tyr
 305 310 315 320
 Leu Gly Val Leu Ala Gly Ser Ser Gly Ala Asn Asp Ala Pro Ala Leu
 325 330 335
 Ile Ser Arg Leu Leu Leu Asp Tyr Arg Glu Ser Pro Ser Leu Ser Leu
 340 345 350
 Leu Glu Asn Val His Gln Val Ala Tyr Glu Ser Asp Gly Thr Ser Cys
 355 360 365

Ala Leu Pro Ala Leu Ala Leu Gly Trp Gln Thr Phe Thr Pro Pro Thr
 370 375 380

Leu Ser Ala Trp Gln Thr Arg Asp Asp Met Gly Lys Leu Ser Leu Leu
 385 390 395 400

Gln Pro Tyr Gln Leu Val Asp Leu Asn Gly Glu Gly Val Val Gly Ile
 405 410 415

Leu Tyr Gln Asp Ser Gly Ala Trp Trp Tyr Arg Glu Pro Val Arg Gln
 420 425 430

Ser Gly Asp Asp Pro Asp Ala Val Thr Trp Gly Ala Ala Ala Ala Leu
 435 440 445

Pro Thr Met Pro Ala Leu His Asn Ser Gly Ile Leu Ala Asp Leu Asn
 450 455 460

Gly Asp Gly Arg Leu Glu Trp Val Val Thr Ala Pro Gly Val Ala Gly
 465 470 475 480

Met Tyr Asp Arg Thr Pro Gly Arg Asp Trp Leu His Phe Thr Pro Leu
 485 490 495

Ser Ala Leu Pro Val Glu Tyr Ala His Pro Lys Ala Val Leu Ala Asp
 500 505 510

Ile Leu Gly Ala Gly Leu Thr Asp Met Val Leu Ile Gly Pro Arg Ser
 515 520 525

Val Arg Leu Tyr Ser Gly Lys Asn Asp Gly Trp Asn Lys Gly Glu Thr
 530 535 540

Val Gln Gln Thr Glu Arg Leu Thr Leu Pro Val Pro Gly Val Asp Pro
 545 550 555 560

Arg Thr Leu Val Ala Phe Ser Asp Met Ala Gly Ser Gly Gln Gln His
 565 570 575

Leu Thr Glu Val Arg Ala Asn Gly Val Arg Tyr Trp Pro Asn Leu Gly
 580 585 590

His Gly Arg Phe Gly Gln Pro Val Asn Ile Pro Gly Phe Ser Gln Ser
 595 600 605

Val Thr Thr Phe Asn Pro Asp Gln Ile Leu Leu Ala Asp Thr Asp Gly
 610 615 620

Ser Gly Thr Thr Asp Leu Ile Tyr Ala Met Ser Asp Arg Leu Val Ile
 625 630 635 640

Tyr Phe Asn Gln Ser Gly Asn Tyr Phe Ala Glu Pro His Thr Leu Leu
 645 650 655

Leu Pro Lys Gly Val Arg Tyr Asp Arg Thr Cys Ser Leu Gln Val Ala
 660 665 670

Asp Ile Gln Gly Leu Gly Val Pro Ser Leu Leu Leu Thr Val Pro His

675	680	685
Val Ala Pro His His Trp	Val Cys His Leu Ser	Ala Asp Lys Pro Trp
690	695	700
Leu Leu Asn Gly Met Asn Asn Asn Met Gly	Ala Arg His Ala Leu His	
705	710	715 720
Tyr Arg Ser Ser Val Gln Phe Trp	Leu Asp Glu Lys Ala Glu Ala Leu	
	725	730 735
Ala Ala Gly Ser Ser Pro Ala Cys Tyr	Leu Pro Phe Thr Leu His Thr	
	740	745 750
Leu Trp Arg Ser Val Val Gln Asp Glu Ile Thr	Gly Asn Arg Leu Val	
	755	760 765
Ser Asp Val Leu Tyr Arg His Gly Val Trp	Asp Gly Gln Glu Arg Glu	
	770	775 780
Phe Arg Gly Phe Gly Phe Val Glu Ile Arg	Asp Thr Asp Thr Leu Ala	
	785	790 795 800
Ser Gln Gly Thr Ala Thr Glu Leu Ser Met	Pro Ser Val Ser Arg Asn	
	805	810 815
Trp Tyr Ala Thr Gly Val Pro Ala Val Asp	Glu Arg Leu Pro Glu Thr	
	820	825 830
Tyr Trp Gln Asn Asp Ala Ala Ala Phe	Ala Asp Phe Ala Thr Arg Phe	
	835	840 845
Thr Val Gly Ser Gly Glu Asp Glu Gln Thr	Tyr Thr Pro Asp Asp Ser	
	850	855 860
Lys Thr Phe Trp Leu Gln Arg Ala Leu Lys	Gly Ile Leu Leu Arg Ser	
	865	870 875 880
Glu Leu Tyr Gly Ala Asp Gly Ser Ser Gln	Ala Asp Ile Pro Tyr Ser	
	885	890 895
Val Thr Glu Ser Arg Pro Gln Val Arg	Leu Val Glu Ala Asn Gly Asp	
	900	905 910
Tyr Pro Val Val Trp Pro Met Gly Ala Glu	Ser Arg Thr Ser Val Tyr	
	915	920 925
Glu Arg Tyr His Asn Asp Pro Gln Cys Gln	Gln Gln Ala Val Leu Leu	
	930	935 940
Ser Asp Glu Tyr Gly Phe Pro Leu Arg Gln	Val Ser Val Asn Tyr Pro	
	945	950 955 960
Arg Arg Pro Pro Ser Ala Asp Asn Pro	Tyr Pro Ala Ser Leu Pro Ala	
	965	970 975
Thr Leu Phe Ala Asn Ser Tyr Asp Glu	Gln Gln Gln Ile Leu Arg Leu	
	980	985 990

Gly Leu Gln Gln Ser Ser Ala His His Leu Val Ser Leu Ser Glu Gly
 995 1000 1005
 His Trp Leu Leu Gly Leu Ala Glu Ala Ser Arg Asp Asp Val Phe Thr
 1010 1015 1020
 Tyr Ser Ala Asp Asn Val Pro Glu Gly Gly Leu Thr Leu Glu His Leu
 025 1030 1035 1040
 Leu Ala Pro Glu Ser Leu Val Ser Asp Ser Gln Val Gly Thr Leu Ala
 1045 1050 1055
 Gly Gln Gln Gln Val Trp Tyr Leu Asp Ser Gln Asp Val Ala Thr Val
 1060 1065 1070
 Ala Ala Pro Pro Leu Pro Pro Lys Val Ala Phe Ile Glu Thr Ala Val
 1075 1080 1085
 Leu Asp Glu Gly Met Val Ser Ser Leu Ala Ala Tyr Ile Val Asp Glu
 1090 1095 1100
 His Leu Glu Gln Ala Gly Tyr Arg Gln Ser Gly Tyr Leu Phe Pro Arg
 105 1110 1115 1120
 Gly Arg Glu Ala Glu Gln Ala Leu Trp Thr Gln Cys Gln Gly Tyr Val
 1125 1130 1135
 Thr Tyr Ala Gly Ala Glu His Phe Trp Leu Pro Leu Ser Phe Arg Asp
 1140 1145 1150
 Ser Met Leu Thr Gly Pro Val Thr Val Thr Arg Asp Ala Tyr Asp Cys
 1155 1160 1165
 Val Ile Thr Gln Trp Gln Asp Ala Ala Gly Ile Val Thr Thr Ala Asp
 1170 1175 1180
 Tyr Asp Trp Arg Phe Leu Thr Pro Val Arg Val Thr Asp Pro Asn Asp
 185 1190 1195 1200
 Asn Leu Gln Ser Val Thr Leu Asp Ala Leu Gly Arg Val Thr Thr Leu
 1205 1210 1215
 Arg Phe Trp Gly Thr Glu Asn Gly Ile Ala Thr Gly Tyr Ser Asp Ala
 1220 1225 1230
 Thr Leu Ser Val Pro Asp Gly Ala Ala Ala Ala Leu Ala Leu Thr Ala
 1235 1240 1245
 Pro Leu Pro Val Ala Gln Cys Leu Val Tyr Val Thr Asp Ser Trp Gly
 1250 1255 1260
 Asp Asp Asp Asn Glu Lys Met Pro Pro His Val Val Val Leu Ala Thr
 265 1270 1275 1280
 Asp Arg Tyr Asp Ser Asp Thr Gly Gln Gln Val Arg Gln Gln Val Thr
 1285 1290 1295

Phe Ser Asp Gly Phe Gly Arg Glu Leu Gln Ser Ala Thr Arg Gln Ala
1300 1305 1310

Glu Gly Asn Ala Trp Gln Arg Gly Arg Asp Gly Lys Leu Val Thr Ala
1315 1320 1325

Ser Asp Gly Leu Pro Val Thr Val Ala Thr Asn Phe Arg Trp Ala Val
1330 1335 1340

Thr Gly Arg Ala Glu Tyr Asp Asn Lys Gly Leu Pro Val Arg Val Tyr
345 1350 1355 1360

Gln Pro Tyr Phe Leu Asp Ser Trp Gln Tyr Val Ser Asp Asp Ser Ala
1365 1370 1375

Arg Gln Asp Leu Tyr Ala Asp Thr His Phe Tyr Asp Pro Thr Ala Arg
1380 1385 1390

Glu Trp Gln Val Ile Thr Ala Lys Gly Glu Arg Arg Gln Val Leu Tyr
1395 1400 1405

Thr Pro Trp Phe Val Val Ser Glu Asp Glu Asn Asp Thr Val Gly Leu
1410 1415 1420

Asn Asp Ala Ser
425

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 973 amino acid residues
- (B) TYPE: amino acid
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: PROTEIN (SepC)

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```

Met Ser Thr Ser Leu Phe Ser Ser Thr Pro Ser Val Ala Val Leu Asp
 1             5             10             15
Asn Arg Gly Leu Leu Val Arg Glu Leu Gln Tyr Tyr Arg His Pro Asp
 20             25             30
Thr Pro Glu Glu Thr Asp Glu Arg Ile Thr Cys His Gln His Asp Glu
 35             40             45
Arg Gly Ser Leu Ser Gln Ser Ala Asp Pro Arg Leu His Ala Ala Gly
 50             55             60
Leu Thr Asn Phe Thr Tyr Leu Asn Ser Leu Thr Gly Thr Val Leu Gln
 65             70             75             80
Ser Val Ser Ala Asp Ala Gly Thr Ser Leu Glu Leu Ser Asp Ala Ala
 85             90             95
Gly Arg Ala Phe Leu Ala Val Thr Gly Ala Gly Thr Glu Asp Ala Val
100             105             110
Thr Arg Thr Trp Gln Tyr Glu Asp Asp Thr Leu Pro Gly Arg Pro Leu
115             120             125
Ser Ile Thr Glu Gln Val Thr Gly Glu Ala Ala Gln Ile Thr Glu Arg
130             135             140
Phe Val Tyr Ala Gly Asn Thr Asp Ala Glu Lys Ile Leu Asn Leu Ala
145             150             155             160
Gly Gln Cys Val Ser His Tyr Asp Thr Ala Gly Leu Val Gln Thr Asp
165             170             175
Ser Ile Ala Leu Ser Gly Val Pro Leu Ala Val Thr Arg Gln Leu Leu
180             185             190
Pro Asp Ala Ala Gly Ala Asn Trp Met Gly Glu Asp Ala Ser Ala Trp
195             200             205
Asn Asp Leu Leu Asp Gly Glu Thr Phe Phe Thr Gln Thr His Ala Asp
210             215             220
Ala Thr Gly Ala Val Leu Ser Ile Thr Asp Ala Lys Gly Asn Leu Gln
225             230             235             240

```

Arg Val Ala Tyr Asp Val Ala Gly Leu Leu Ser Gly Ser Trp Leu Thr
 245 250 255
 Leu Lys Asp Gly Thr Glu Gln Val Ile Val Ala Ser Leu Thr Tyr Ser
 260 265 270
 Ala Ala Gly Lys Lys Leu Arg Glu Glu His Gly Asn Gly Val Val Thr
 275 280 285
 Ser Tyr Ile Tyr Glu Pro Glu Thr Gln Arg Leu Thr Gly Ile Lys Thr
 290 295 300
 Glu Arg Pro Ser Gly His Val Ala Gly Ala Lys Val Leu Gln Asp Leu
 305 310 315 320
 Arg Tyr Thr Tyr Asp Pro Val Gly Asn Val Leu Ser Val Asn Asn Asp
 325 330 335
 Ala Glu Glu Thr Arg Phe Trp Arg Asn Gln Lys Val Val Pro Glu Asn
 340 345 350
 Thr Tyr Ile Tyr Asp Ser Leu Tyr Gln Leu Val Ser Ala Thr Gly Arg
 355 360 365
 Glu Met Ala Asn Ala Gly Gln Gln Gly Asn Asp Leu Pro Ser Ala Thr
 370 375 380
 Ala Pro Leu Pro Thr Asp Ser Ser Ala Tyr Thr Asn Tyr Thr Arg Thr
 385 390 395 400
 Tyr Arg Tyr Asp Arg Gly Gly Asn Leu Thr Gln Met Arg His Ser Ala
 405 410 415
 Pro Ala Thr Asn Asn Asn Tyr Thr Thr Asp Ile Thr Val Ser Asp Arg
 420 425 430
 Ser Asn Arg Ala Val Leu Ser Thr Leu Ala Glu Val Pro Ser Asp Val
 435 440 445
 Asp Met Leu Phe Ser Ala Gly Gly His Gln Lys His Leu Gln Pro Gly
 450 455 460
 Gln Ala Leu Val Trp Thr Pro Arg Gly Glu Leu Gln Lys Val Thr Pro
 465 470 475 480
 Val Val Arg Asp Gly Gly Ala Asp Asp Ser Glu Ser Tyr Arg Tyr Asp
 485 490 495
 Ala Gly Ser Gln Arg Ile Ile Lys Thr Gly Thr Arg Gln Thr Gly Asn
 500 505 510
 Asn Val Gln Thr Gln Arg Val Val Tyr Leu Pro Gly Leu Glu Leu Arg
 515 520 525
 Ile Met Ala Asn Gly Val Thr Glu Lys Glu Ser Leu Gln Val Ile Thr
 530 535 540
 Val Gly Glu Ala Gly Arg Ala Gln Val Arg Val Leu His Trp Glu Ile

545 550 555 560
 Gly Lys Pro Asp Asp Leu Asp Glu Asp Ser Val Arg Tyr Ser Tyr Asp
 565 570 575
 Asn Leu Val Gly Ser Ser Gln Leu Glu Leu Asp Arg Glu Gly Tyr Leu
 580 585 590
 Ile Ser Glu Glu Glu Phe Tyr Pro Tyr Gly Gly Thr Ala Val Leu Thr
 595 600 605
 Ala Arg Ser Glu Val Glu Ala Asp Tyr Lys Thr Ile Arg Tyr Ser Gly
 610 615 620
 Lys Glu Arg Asp Ala Thr Gly Leu Asp Tyr Tyr Gly Tyr Arg Tyr Tyr
 625 630 635 640
 Gln Pro Trp Ala Gly Arg Trp Leu Ser Thr Asp Pro Ala Gly Thr Val
 645 650 655
 Asp Gly Leu Asn Leu Phe Arg Met Val Arg Asn Asn Pro Val Thr Leu
 660 665 670
 Phe Asp Ser Asn Gly Arg Ile Ser Thr Gly Gln Glu Ala Arg Arg Leu
 675 680 685
 Val Gly Glu Ala Phe Val His Pro Leu His Met Pro Val Phe Glu Arg
 690 695 700
 Ile Ser Val Glu Arg Lys Ile Ser Met Ser Val Arg Glu Ala Gly Ile
 705 710 715 720
 Tyr Thr Ile Ser Ala Leu Gly Glu Gly Ala Ala Ala Lys Gly His Asn
 725 730 735
 Ile Leu Glu Lys Thr Ile Lys Pro Gly Ser Leu Lys Ala Ile Tyr Gly
 740 745 750
 Asp Lys Ala Glu Ser Ile Leu Gly Leu Ala Lys Arg Ser Gly Leu Val
 755 760 765
 Gly Arg Val Gly Gln Trp Asp Ala Ser Gly Val Arg Gly Ile Tyr Ala
 770 775 780
 His Asn Arg Pro Gly Gly Glu Asp Leu Val Tyr Pro Val Ser Leu Gln
 785 790 795 800
 Asn Thr Ser Ala Asn Glu Ile Val Asn Ala Trp Ile Lys Phe Lys Ile
 805 810 815
 Ile Thr Pro Tyr Thr Gly Asp Tyr Asp Met His Asp Ile Ile Lys Phe
 820 825 830
 Ser Asp Gly Lys Gly His Val Pro Thr Ala Glu Ser Ser Glu Glu Arg
 835 840 845
 Gly Val Lys Asp Leu Ile Asn Lys Gly Val Ala Glu Val Asp Pro Ser
 850 855 860

Arg Pro Phe Glu Tyr Thr Ala Met Asn Val Ile Arg His Gly Pro Gln
865 870 875 880

Val Asn Phe Val Pro Tyr Met Trp Glu His Glu His Asp Lys Val Val
885 890 895

Asn Asp Asn Gly Tyr Leu Gly Val Val Ala Ser Pro Gly Pro Phe Pro
900 905 910

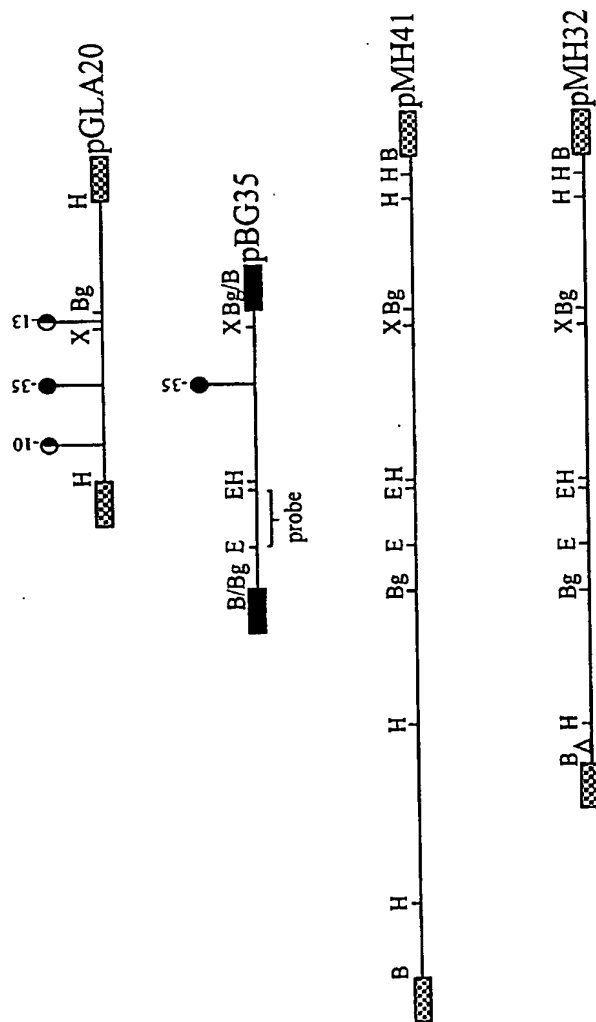
Val Ala Met Val His Gln Gly Glu Trp Thr Val Phe Asp Asn Ser Glu
915 920 925

Glu Leu Phe Asn Phe Tyr Lys Ser Thr Asn Thr Pro Leu Pro Glu His
930 935 940

Trp Ser Gln Asp Phe Met Asp Arg Gly Lys Gly Ile Val Ala Thr Pro
945 950 955 960

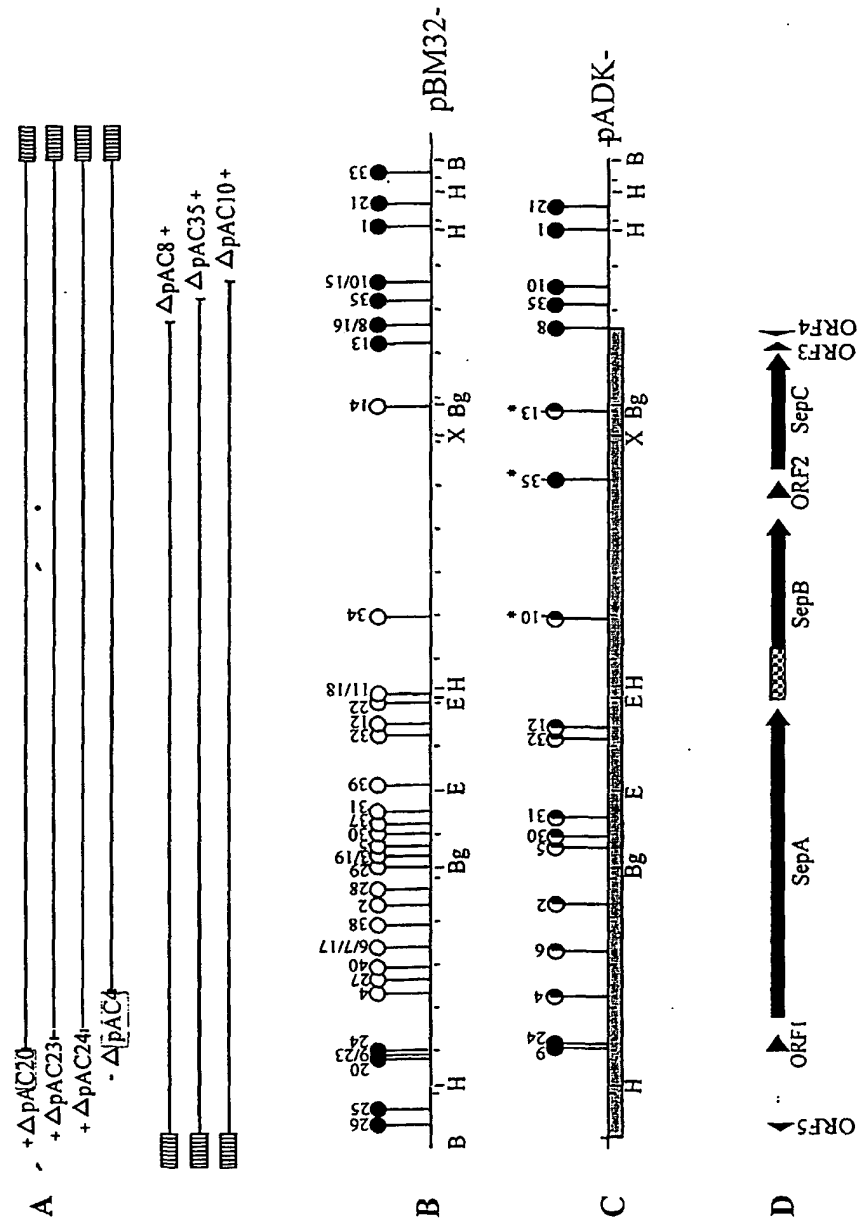
Arg His Ala Glu Leu Leu Asp Lys Arg Arg Val Met Tyr
965 970

FIGURE 1



WO 01/16305

PCT/NZ00/00174

FIGURE 2

WO 01/16305

PCT/NZ00/00174

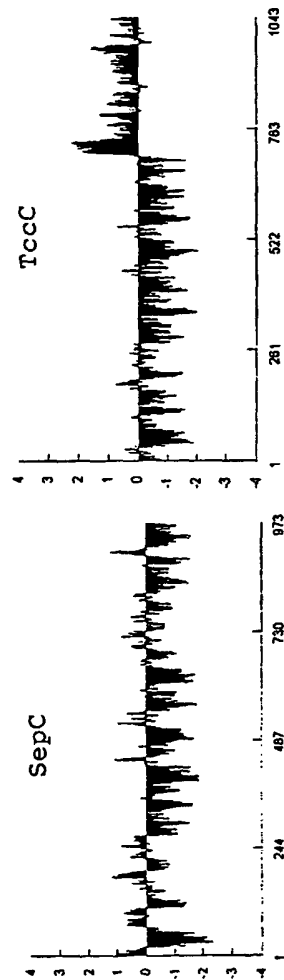
FIGURE 3

FIGURE 4 - continued

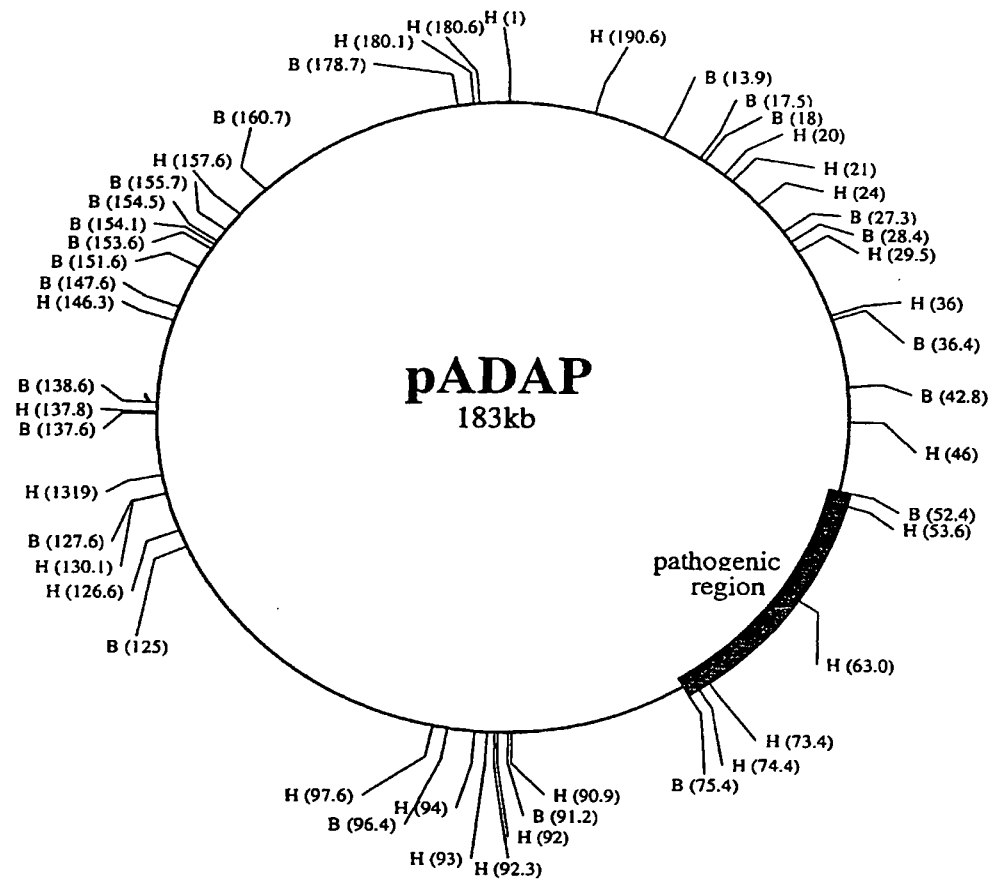
Sepa	1251	PPESDLVCSPPSRVAVAVEN --- SOUTEASTTITCOSEIUKISPRVDSSESSSEIICNAPPETOCNVTNPLTECAGCFTS...VJSKEDDIAQ.I
Toda	1255	UPKSTIERITKSRMLALIGOKITKTCGSGOGVOEDPNEMLAHKRVDSWHEIFEDNKKSEKILANEP --- EYACAGOGODLAWFACQDT
TcBa	1235	UPKANTIRPVPVSRMLALIG --- SOUTEASTTITCOSEIUKISPRVDSSESSSEIICNAPPETOCNVTNPLTECAGCFTS...VJSKEDDIAQ.I
TcCb	274	SEANGGVEPPWE --- NK --- IIPFTISKEDKIDVWVWVWSSDSWAK --- IIPFTISKEDKIDVWVWVWSSDSWAK --- IIPFTISKEDKIDVWVWVWSSDSWAK
TcCb	334	TVLEHVRVENDRIVAVVERD --- PAVGDDGCGIKGTHAHTIPOTKRYDHTIPATITLATOOGESSEIIPESLLID --- ESSTTLQVWVWVWSSDSWAK
Sepa	1334	SVNDMTLNVISDIGIKSKEFEN --- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
Toda	1327	LDSEVN --- RSMOEXIPEADAGQUHPEQ --- SAVPEISYQODPNAV --- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcBa	1320	YSSTITVAPVETATPEADAGQUHPEQ --- ATNPNSTPOTPVADPPSPKVVITRWVWVWSSDSWAK --- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcCb	295	TEETDSPPYCRILAGVPRQFECGAGNKKPVVGYLYCTSAFNGVLR --- SKNPLSTITPTEGCGSLOQVPOKQVTKVTCATEPPE --- ETEKEMVDSDFG
Sepa	1423	HIFANVSCFSPADITCTGSLTPVOAGIEPPEIPEVAG --- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
Toda	1435	YSESANV --- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcBa	1428	STRVJESVMPITVESADIPUSNALSITINGVAGTHPCHILKAGVSGKPEHDS --- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcCb	361	ETEVASITGSAONIBDGTVLIFQAG --- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcCb	528	VDDLKGTGAYVYV --- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
Sepa	1510	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
Toda	1532	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcBa	1525	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcCb	637	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
Sepa	1588	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
Toda	1603	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcBa	1597	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcCb	409	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcCb	744	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
Sepa	1620	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
Toda	1711	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcBa	1707	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcCb	441	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcCb	783	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
Sepa	1687	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
Toda	1815	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcBa	1810	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcCb	509	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcCb	845	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
Sepa	1796	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
Toda	1924	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcBa	1919	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcCb	606	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL
TcCb	955	--- ETEKEMVDSDFG --- AVEKESGLIDVTPSPILLNSKNTSTVFTNEDSSILTPRL

FIGURE 4 - continued

SepA	1888
Tcda	2019
Tcda	2013
Tcda	712
Tcda	1055
SepA	1998
Tcda	2129
Tcda	2123
Tcda	822
Tcda	1165
SepA	2108
Tcda	2239
Tcda	2233
Tcda	932
Tcda	1275
SepA	2215
Tcda	2348
Tcda	2341
Tcda	1042
Tcda	1380
SepA	2300
Tcda	2438
Tcda	2429
Tcda	1117
Tcda	1455
SepA	2376
Tcda	2316
Tcda	2304
Tcda	1189
Tcda	1565

FIGURE 5

SepC 101 ASFG--LFSSTLSAVLDRGHLVRELQYGHDPYPTDSEHLCGQDQCSLSADPRPLAG-----LNFETLNSAGCTTOSVAGDNGTSLDSDAGRAFL
 Tccc 109 KSPSTLITVTHVSLDRGLSLSDIGHFIVIG-GDITVTFEGIDKGLHNSIOPEHTEKQADNSVKEHVMCHDLAEPARTSVDRGTVADGTEGSMV
 SepC 210 AVTCASTEDATVTRMOCEDTILPGHPLSTEQVIG-SAQOTIEFPYAGTNDIENITILILAGQVSRVDITAGLVOTISIALSEVAVTITROLPPDAAGANRGGDSAND
 Tccc 215 TMAVAG----VRCIPRMEGNTLPGHPLSVSQFNCEKVTERTKAGNTISENIEYHLSGLITVHTDAGVTRLSSSLANLQSQHLLARGQENRSGDETVMCG
 SepC 320 LDCETPTQTHADTISANLSTIDAKCHLORAVDVGCLSSKSLITLCCTEQUVAVSTTSNAGCKLREKNGCVTSTTEPETORTGHTDPSQVAGVQVLODI
 Tccc 325 MLCSEVITLCTSTTALISNLTCTIDAKCHLORAVDVGCLSSKSLITLCCTEQUVAVSTTSNAGCKLREKNGCVTSTTEPETORTGHTDPSQVAGVQVLODI
 SepC 430 RYTHDPVGNWLSUNDAETREWRHQRVWEHTTHYDSLVLVSATGREMAMAGQCNLPSATAELPTDESAVTNYRTLYDRGCHLQRIISAPATMNTTIDITVS
 Tccc 435 RYTHDPVGNWLSUNDAETREWRHQRVWEHTTHYDSLVLVSATGREMAMAGQCNLPSATAELPTDESAVTNYRTLYDRGCHLQRIISAPATMNTTIDITVS
 SepC 540 DRSRAVLSTLAVQSDVDLFSAGCHORLCPQALHTFAGELQVTPVTDGADSESTEDAGGSCILITCTGTGTGNTVQRTVILPGLHPLMANGVTETESL
 Tccc 544 DRSRAVLSTLHTTITVDALEDSGCHORLTPQCNLNTIRGELQVTPVDSNSQSEWTRISDCKGLLVSSQDIGHSTVQRTVILPGLHPLMANGVTETESL
 SepC 650 QVITVGEAGRAQVRVLHREISKPLDLEDSVRYSYDILAGSSQLELDEQCLLSDEEPYVPGTAVLTAISEVADYKTRYSKGRDNTGLTYGVRYTQPIAGRNLSG
 Tccc 654 QVITVGEAGRAQVRVLHREISKPLDLEDSVRYSYDILAGSSQLELDEQCLLSDEEPYVPGTAVLTAISEVADYKTRYSKGRDNTGLTYGVRYTQPIAGRNLSG
 SepC 735 DPAGTVGGLMLRWVRNRPVTLFENGRISTQQRRLVGSFVPELHMPVTERISVERKISNVRDSGTYTTSAGDMAG-----H
 Tccc 764 DPAGTVGGLMLRWVRNRPVTLFENGRISTQQRRLVGSFVPELHMPVTERISVERKISNVRDSGTYTTSAGDMAG-----H
 SepC 818 NILEKT-----IKP-----G--SLAKHGDQDESILGLAKRGLVSGCQDASGVGIVAHNRFGSD-----LVTPVSLQNTSANEIVNPKFKIIL
 Tccc 874 NILEKT-----IKP-----G--SLAKHGDQDESILGLAKRGLVSGCQDASGVGIVAHNRFGSD-----LVTPVSLQNTSANEIVNPKFKIIL
 SepC 926 PYTGDDYHDIKPSDC-KGHVPTASSEERAVDLINSAVEDDSRPFFYTNQVTRGQVNFVPTVMHEHDKVNDNGTLGVMSPPFVAHQCEKTVFDN-
 Tccc 926 QGSTRAGIACITTYSWIGFGLDVAISNPAHLANYAVGTAAGLG----ALEKAVRQCGSFLSRLLRVSVVPAAGLARQLVHPSVRRVPEIFPSLGLGLVCGIGTC
 SepC 973 -----SGLFNFTYGTNTPELPEMVSODPMORGHTLAPR-LVELLX-----RWVY- 973
 Tccc 1043 LHRVWGRSWMISBALSAAGSGIDVACHIGNQIHEPLTTGILNALVGTSAVGNRDEPSL 1043

FIGURE 6

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Glare, Travis T
Hurst, Mark R H
Jackson, Trevor A
- (ii) TITLE OF INVENTION: Insecticidal nucleotide sequences
- (iii) NUMBER OF SEQUENCES: 6
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: A J Park & Son
(B) STREET: Huddart Parker Building, Post Office Square
(C) CITY: Wellington
(D) COUNTRY: New Zealand
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:

(2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18937 nucleotides (A) LENGTH: 5118 amino acids
(B) TYPE: nucleotide (B) TYPE: amino acid
(C) STRANDEDNESS: single (C) STRANDEDNESS:
(D) TOPOLOGY: Linear (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA (ii) MOLECULE TYPE: PROTEIN
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ggatccgagt gaaggaatca teggccgctt tatacgtttc aggggaata cgggtggccg 60
caacgtggca atggatgttg tttgtgtcgg tatgaatcgc cgcaacgtac tgggtgttcg 120
acatacccag tgccgataaa ctgtgacgaa cactatcaaa gatgtgttcc gtcgacctga 180
aagccaggat ttatctttac accaatgggt ggggtgggctt cctttctgaa ctgggtgcac 240
atctagcgg catcatcaaa agatgcatgg aaatacaaat atcatattta cagacacca 300
agttgatgac ctgctccgtg agttgaaatg ccgacggggg aaatcagcag ccttttcaac 360
tcatggagca gggggaaatc aatcctcaat aaccgcatt ggatatcctg ccagtgtgca 420
tttaaccttt ttagtgtgtt tccttaatat cccaatcgtt gaatcgctac atacggcaga 480
cattagtatc tcacttatca tcaaagtaat atcacaccga gaatgctaatt tcatgatgat 540
gaaaacgttc cattaataaa ttttcagaaa cctaacacgg catttttatg ctgatcagt 600
aattgattgt ttctgaaaaa attaattgca cctctgccac ttatcagata aaaacacccc 660
atgcggtaag ttttttattt tttattaatg attttattaa tgattttatt aatgatttta 720
ttaatgattt tattaatgat ttactatag atgaatgtta acatgggtga taatttactt 780
tactcaattt aattgttggg atgaccatgt ttagatgag tggcacggat tcattattgt 840
aaaaaaagta tctaaaacct ttagcagcaa tcctacttga ggatgacctc gacaggactt 900
gattattgcc attttttacg aaggaagatg acgggtgata aataataaaa aaaacaaaag 960
tatagcctta ggtatcgccg attacatcca gtaacactta ttgacttttt ttactttcta 1020

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
8 March 2001 (08.03.2001)

PCT

(10) International Publication Number
WO 01/16305 A3

- (51) International Patent Classification⁷: C12N 15/31, 15/70, 15/82, C07K 14/24, C12Q 1/68, A01N 63/02, A01H 5/00
- (74) Agent: WILSON, Kathryn, S.; all of Level 12., KPMG Center., 85 Alexandra Street, Private Bag 3140, Hamilton (NZ).
- (21) International Application Number: PCT/NZ00/00174
- (22) International Filing Date:
4 September 2000 (04.09.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
337610 2 September 1999 (02.09.1999) NZ
- (71) Applicant (for all designated States except US): AGRE-SEARCH LIMITED [NZ/NZ]; 5th floor, Tower Block, Ruakura Research Centre, East Street, Hamilton 2001 (NZ).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): GLARE, Travis, Robert [AU/NZ]; 38 Whincorps Road, Halswell, Christchurch 8003 (NZ). HURST, Mark, Robin, Holmes [NZ/NZ]; 148 Hendersons Road, Hoon Hay, Christchurch 8002 (NZ). JACKSON, Trevor, Anthony [NZ/NZ]; 407 Halswell Road, Halswell, Christchurch 8003 (NZ).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:
— with international search report
- (88) Date of publication of the international search report:
10 January 2002
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/16305 A3

(54) Title: NUCLEOTIDE SEQUENCES ENCODING AN INSECTIDAL PROTEIN COMPLEX FROM SERRATIA

(57) Abstract: The present invention concerns novel nucleotide sequences encoding proteins from the Enterobacteriaceae, *Serratia entomophila* and *Serratia proteamaculans*, and the use of said nucleotide sequences and proteins for inherent insecticidal and potentially metazoacidal properties. The invention relates to an isolated nucleic acid molecule comprising a nucleotide sequence that encodes an insecticidal protein complex, or a functional fragment, neutral mutation, or homolog thereof capable of hybridising with the nucleic acid molecule under standard hybridisation conditions. The nucleotide sequences include a pathogenicity-encoding region cloned from bacteria *Serratia entomophila* and *S. proteamaculans*. The region contain pathogenic determinants of a disease that affect the grass grub, *Costelytra zealandica* Coleoptera: Scarabaeidae, an important insect pasture pest in New Zealand. The proteins encoded by determined genes may be used for insect control whether as an inundative pesticide, within baits or expressed in other organisms such as plants or microbes.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/NZ 00/00174

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/31 C12N15/70 C12N15/82 C07K14/24 C12Q1/68
A01N63/02 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K A01H C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

BIOSIS, EPO-Internal, STRAND, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
------------	--	-----------------------

X	<p>JACKSON T A ET AL: "PATHOGEN TO PRODUCT DEVELOPMENT OF SERRATIA-ENTOMOPHILA ENTEROBACTERIACEAE AS A COMMERCIAL BIOLOGICAL CONTROL AGENT FOR NEW ZEALAND GRASS GRUB COSTELYTRA-ZEALANDICA" JACKSON, T. A. AND T. R. GLARE (ED.). USE OF PATHOGENS IN SCARAB PEST, 1992, pages 191-198, XP000997900 0-946707-35-9. 1992 the whole document</p> <p style="text-align: center;">---</p> <p style="text-align: center;">-/--</p>	32
---	--	----

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

23 May 2001

Date of mailing of the international search report

06/06/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Holtorf, S

INTERNATIONAL SEARCH REPORT

International Application No

PCT/NZ 00/00174

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>GRKOVIC STEVE ET AL: "Genes Essential for Amber Disease in Grass Grubs Are Located on the Large Plasmid Found in Serratia entomophila and Serratia proteamaculans." APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 61, no. 6, 1995, pages 2218-2223, XP000994573 ISSN: 0099-2240 cited in the application the whole document</p> <p>---</p>	
A	<p>GLARE TRAVIS R ET AL: "Plasmid transfer among several members of the family Enterobacteriaceae increases the number of species capable of causing experimental amber disease in grass grub." FEMS MICROBIOLOGY LETTERS, vol. 139, no. 2-3, 1996, pages 117-120, XP000998482 ISSN: 0378-1097 cited in the application the whole document</p> <p>---</p>	
A	<p>WO 99 42589 A (NOVARTIS ERFIND VERWALT GMBH ;NOVARTIS AG (CH); KRAMER VANCE CARY) 26 August 1999 (1999-08-26) the whole document</p> <p>---</p>	
A	<p>WO 98 08932 A (DOW AGROSCIENCES LLC ;WISCONSIN ALUMNI RES FOUND (US)) 5 March 1998 (1998-03-05) the whole document</p> <p>---</p>	
A	<p>WO 98 08388 A (MORGAN JAMES ALUN WYNNE ;JARRETT PAUL (GB); ELLIS DEBORAH JUNE (GB) 5 March 1998 (1998-03-05) the whole document</p> <p>---</p>	
A	<p>WO 97 17432 A (WISCONSIN ALUMNI RES FOUND) 15 May 1997 (1997-05-15) the whole document</p> <p>---</p>	
A	<p>BOWEN D ET AL: "INSECTICIDAL TOXINS FROM THE BACTERIUM Photorhabdus limnescens" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, US, vol. 280, 26 June 1998 (1998-06-26), pages 2129-2132, XP002115650 ISSN: 0036-8075 cited in the application</p> <p>---</p> <p style="text-align: center;">-/--</p>	

INTERNATIONAL SEARCH REPORT

International Application No

PCT/NZ 00/00174

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>NUNEZ-VALDEZ M E ET AL: "The amb2 locus from <i>Serratia entomophila</i> confers anti-feeding effect on larvae of <i>Costelytra zealandica</i> (Coleoptera: Scarabaeidae)"</p> <p>GENE,NL,ELSEVIER BIOMEDICAL PRESS. AMSTERDAM, vol. 172, no. 1, 12 June 1996 (1996-06-12), pages 75-79, XP004042712</p> <p>ISSN: 0378-1119</p> <p>cited in the application</p>	
P,X	<p>HURST MARK R H ET AL: "Plasmid-located pathogenicity determinants of <i>Serratia entomophila</i>, the causal agent of amber disease of grass grub, show similarity to the insecticidal toxins of <i>Photobacterium luminescens</i>."</p> <p>JOURNAL OF BACTERIOLOGY, vol. 182, no. 18, September 2000 (2000-09), pages 5127-5138, XP002166799</p> <p>ISSN: 0021-9193</p> <p>the whole document</p>	<p>1-4, 9-16, 21-27, 31,41</p>

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 17

Present claim 17 relates to a ligand defined by reference to a desirable characteristic or property, namely binding to the polypeptide of claim 15.

The claim covers all products having this characteristic or property, whereas the application provides no support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for such products.

In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product by reference to a result to be achieved.

Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/NZ 00/00174

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9942589	A	26-08-1999	AU 3028699 A EP 1054972 A	06-09-1999 29-11-2000
WO 9808932	A	05-03-1998	AU 729228 B AU 1050997 A AU 2829997 A BR 9606889 A BR 9711441 A CA 2209659 A EP 0797659 A EP 0970185 A HU 9900768 A JP 2000515024 T PL 321212 A PL 332033 A SK 24699 A SK 93197 A TR 9901126 T WO 9717432 A	25-01-2001 29-05-1997 19-03-1998 28-10-1997 24-10-2000 15-05-1997 01-10-1997 12-01-2000 28-06-1999 14-11-2000 24-11-1997 16-08-1999 10-04-2000 06-05-1998 21-07-1999 15-05-1997
WO 9808388	A	05-03-1998	AU 4024997 A BR 9711285 A CN 1233938 A EP 0923295 A TR 9900435 T ZA 9707373 A	19-03-1998 17-08-1999 03-11-1999 23-06-1999 21-06-1999 15-02-1999
WO 9717432	A	15-05-1997	AU 729228 B AU 1050997 A BR 9606889 A CA 2209659 A EP 0797659 A HU 9900768 A PL 321212 A PL 332033 A SK 93197 A AU 2829997 A BR 9711441 A EP 0970185 A JP 2000515024 T SK 24699 A TR 9901126 T WO 9808932 A	25-01-2001 29-05-1997 28-10-1997 15-05-1997 01-10-1997 28-06-1999 24-11-1997 16-08-1999 06-05-1998 19-03-1998 24-10-2000 12-01-2000 14-11-2000 10-04-2000 21-07-1999 05-03-1998

common function.

The present applicant has now found that three regions of the pADAP plasmid are required for full insecticidal function. Sequence analysis of these three regions has shown that the present applicant has isolated and identified a novel toxin from *Serratia* species that belongs to a new family of insecticidal toxins. It is broadly to this toxin that the present invention is directed.

DISCLOSURE OF INVENTION

According to a first aspect of the present invention, there is provided an isolated nucleic acid molecule comprising a nucleotide sequence of SEQ ID NO: 1 which encodes an insecticidal protein complex, or a functional fragment, neutral mutation, or homolog thereof capable of hybridising with said nucleic acid molecule under standard hybridisation conditions.

The invention also provides an isolated nucleic acid molecule comprising the nucleotide sequence 1995-18937 of SEQ ID NO: 1 which encodes an insecticidal protein complex, or a functional fragment, neutral mutation, or homolog thereof capable of hybridising with said nucleic acid molecule under standard hybridisation conditions.

The invention also provides an isolated nucleic acid molecule comprising one or more of the nucleotide sequences 2411-9547, 9589-13883 or 14546-17467 of SEQ ID NO: 1 which encode insecticidal proteins, or a functional fragment, neutral mutation, or homolog thereof capable of hybridising with said nucleic acid molecule under standard hybridisation conditions.

Preferably the nucleic acid molecule comprises all of nucleotide sequences 2411-9547, 9598-13884 and 14546-17467 of SEQ ID NO: 1.

The invention further relates to an isolated nucleic acid molecule comprising a sequence of SEQ ID NO: 1, nucleotides 1955-18937 of SEQ ID NO: 1 or one or more of nucleotides 2411-9547, 9598-13884 or 14546-17467 of SEQ ID NO: 1, operably linked to at least one further nucleotide sequence which encode an insecticidal protein. For example, the at least one further nucleotide sequence may be the nucleotide sequence which codes for the *Bacillus delta* endo toxins, vegetative insecticidal proteins (vips), cholesterol oxidases, *Clostridium bifermentens* mosquitocidal toxins and/or *Photorhabdus luminescens* toxins and so forth.

The nucleic acid molecule may comprise DNA, cDNA or RNA.

- 10 Preferably said fragment, neutral mutation or homolog thereof is capable of hybridising to said nucleic acid molecule under stringent hybridisation conditions.

The invention further relates to nucleic acid molecules which hybridise to the nucleotide sequence of SEQ ID NO: 1, or nucleotides 1955-18937, 2411-9547, 9598-13884 or 14546-17467 of SEQ ID NO: 1 if there is at least 50%, preferably 60%, more preferably 70% and most preferably 90-95% or greater identity between the sequences.

The nucleic acid molecule may be isolated from *Serratia entomophila* or *Serratia proteamaculans* strains.

Also provided by the present invention are recombinant expression vectors containing the nucleic acid molecule of the invention and hosts transformed with the vector of the invention capable of expressing a polypeptide of the invention.

The vector may be selected from any suitable natural or artificial plasmid/vector. For example, pUC 19 (Yannish-Perron et al. 1995), pProEX HT (GibcoBRL, Gaithersburg, MD, USA), pBR322 (Bolivar et al. 1977), pACYC184 (Chang et al. 1978), pLAFR3 (Staskowicz et al. 1987), and so forth.

THE CLAIMS DEFINING THE INVENTION ARE:

1. A purified and isolated nucleic acid molecule comprising a nucleotide sequence of SEQ ID NO: 1 that encodes at least one of:
 - (i) an insecticidal protein complex, or
 - (ii) a functional fragment of said complex, or
 - (iii) a neutral mutation of said complex, or
 - (iv) a homolog of said complex,each of which are capable of hybridising with said nucleic acid molecule under standard hybridisation conditions.
2. A purified and isolated nucleic acid molecule as claimed in Claim 1 comprising the nucleotide sequence 1995-18937 of SEQ ID NO: 1.
3. A purified and isolated nucleic acid molecule as claimed in Claim 1 comprising one or more of the nucleotide sequences 2411-9547, 9589-13883 or 14546-17467 of SEQ ID NO: 1.
4. A purified and isolated nucleic acid molecule as claimed in Claim 3 comprising all of nucleotide sequences 2411-9547, 9598-13884 and 14546-17467 of SEQ ID NO: 1.
5. A purified and isolated nucleic acid molecule as claimed in Claim 1 comprising a sequence of SEQ ID NO: 1, operably linked to at least one further nucleotide sequence which encode an insecticidal protein.
6. A purified and isolated nucleic acid molecule as claimed in Claim 2 comprising nucleotides 1955-18937 of SEQ ID NO: 1, operably linked to at least one further nucleotide sequence which encode an insecticidal protein.

7. A purified and isolated nucleic acid molecule as claimed in Claim 3 comprising a sequence of SEQ ID NO: 1, or one or more of nucleotides 2411-9547, 9598-13884 or 14546-17467 of SEQ ID NO: 1, operably linked to at least one further nucleotide sequence which encode an insecticidal protein.
8. A purified and isolated nucleic acid molecule as claimed in any one of claims 4 through 6 wherein the said nucleotide sequence includes the nucleotide sequence which codes for at least one of the *Bacillus delta endo* toxins, vegetative insecticidal proteins (vips), cholesterol oxidases, *Clostridium bifermentens* mosquitocidal toxins and/or *Photobacterium luminescens* toxins.
9. A purified and isolated nucleic acid molecule as claimed in claim 1 wherein nucleic acid molecule may comprise DNA, cDNA or RNA.
10. A purified and isolated nucleic acid molecule as claimed in claim 1 wherein the nucleic acid molecules said fragment, neutral mutation or homolog thereof capable of hybridising to said nucleic acid molecule, hybridise to the nucleotide sequence of SEQ ID NO: 1, or nucleotides 1955-18937, 2411-9547, 9598-13884 or 14546-17467 of SEQ ID NO: 1 if there is at least 50%, preferably 60%, more preferably 70% and most preferably 90-95% or greater identity between the sequences.
11. A purified and isolated nucleic acid molecule as claimed in claim 1 wherein the nucleic acid molecule may be isolated from *Serratia entomophila* or *Serratia proteamaculans* strains of bacteria.
12. A recombinant expression vector(s) containing the nucleic acid molecule as claimed in Claim 1 and host transformed with the vector expressing a polypeptide.
13. A recombinant expression vector(s) as claimed in claim 11 wherein the vector is selectable from any suitable natural or artificial plasmid/vector.
14. A recombinant expression vector(s) as claimed in claim 13 wherein said suitable natural or

artificial plasmid/vector, including, pUC 19 (Yannish-Perron et al. 1995), pProEX HT (GibcoBRL, Gaithersburg, MD, USA), pBR322 (Bolivar et al. 1977), pACYC184 (Chang et al. 1978), pLAFR3 (Staskowicz et al. 1987).

15. A polypeptide resulting from the transformation or transfection of a host cell with a recombinant expression vector as claimed in any one of Claims 12 through 14.
16. A method of producing a polypeptide of claim 15 comprising the steps of:
 - (a) culturing a host cell which has been transformed or transfected with said vector as defined above to express the encoded polypeptide or peptide; and
 - (b) recovering the expressed polypeptide or peptide.
17. A ligand that binds to a polypeptide of Claim 15.
18. A ligand as claimed in claim 17 wherein the ligand is an antibody or antibody binding fragment.
19. Probes and primers comprising a fragment of the nucleic acid molecule as claimed in Claim 1 wherein said fragment is hybridisable under stringent conditions to a native insecticidal gene sequence.
20. Probes and primers comprising a fragment of the nucleic acid molecule as claimed in claim 19 wherein said probes and primers enable the structure and function of the gene to be determined and homologs of the gene to be obtained from bacteria other than *Serratia* sp.
21. A polypeptide as claimed in Claim 15 wherein the polypeptide has insecticidal activity encoded by the nucleic acid molecule of claim 1, or a functional fragment, neutral mutation or homolog thereof.
22. A polypeptide having insecticidal activity as claimed in claim 21 wherein the polypeptide